8-2017

# Effective Visualizations of the Uncertainty in Hurricane Forecasts

Le Liu
*Clemson University*, lel@g.clemson.edu

# EFFECTIVE VISUALIZATIONS OF THE UNCERTAINTY IN HURRICANE FORECASTS

---

A Dissertation
Presented to
the Graduate School of
Clemson University

---

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Computer Science

---

by
Le Liu
August 2017

---

Accepted by:
Dr. Donald H. House, Committee Chair
Dr. Joshua A. Levine
Dr. Brian C. Dean
Dr. Amy W. Apon

# Abstract

The track forecast cone developed by the U.S. National Hurricane Center is the one most universally adopted by the general public, the news media, and governmental officials to enhance viewers' understanding of the forecasts and their underlying uncertainties. However, current research has experimentally shown that it has limitations that result in misconceptions of the uncertainty included. Most importantly, the area covered by the cone tends to be misinterpreted as the region affected by the hurricane. In addition, the cone summarizes forecasts for the next three days into a single representation and, thus, makes it difficult for viewers to accurately determine crucial time-specific information.

To address these limitations, this research develops novel alternative visualizations. It begins by developing a technique that generates and smoothly interpolates robust statistics from ensembles of hurricane predictions, thus creating visualizations that inherently include the spatial uncertainty by displaying three levels of positional storm strike risk at a specific point in time. To address the misconception of the area covered by the cone, this research develops time-specific visualizations depicting spatial information based on a sampling technique that selects a small, representative subset from an ensemble of points. It also allows depictions of such important storm characteristics as size and intensity. Further, this research generalizes the representative sampling framework to process ensembles of forecast tracks, selecting a subset of tracks accurately preserving the original distributions of available storm characteristics and keeping appropriately defined spatial separations. This framework supports an additional hurricane visualization portraying prediction uncertainties implicitly by directly showing the members of the subset without the visual clutter. We collaborated on cognitive studies that suggest that these visualizations enhance viewers' ability to understand the forecasts because they are potentially interpreted more like uncertainty distributions.

In addition to benefiting the field of hurricane forecasting, this research potentially enhances the visualization community more generally. For instance, the representative sampling framework for processing 2D points developed here can be applied to enhancing the standard scatter plots and density plots by reducing

sizes of data sets. Further, as the idea of direct ensemble displays can possibly be extended to more general numerical simulations, it, thus, has potential impacts on a wide range of ensemble visualizations.

# Dedication

This dissertation is dedicated to my grandfather XiaoKai Zhang, who passed away during my pursuit of this PhD Degree. His eagerness to learn and his spirit of never giving up kept encouraging me to complete my degree. This dissertation is also dedicated to my grandmother Xia Gao, my parents Longsheng Liu and Jingli Zhang, my parents-in-law Heping Yang and Huimin Liu, and my wife Xu Yang. Their selfless love, patience, and encouragement were always my strongest support throughout this challenging journey.

# Acknowledgments

I would like to express my sincere appreciations to several people who helped me in the completion of this dissertation. Most importantly, I would like to thank my supervisor, Dr. Donald House, for his Computer Graphics class, which led me to this interesting field, and for his patient, wise, and professional guidance during the past six years. His mentoring was my biggest help throughout this research.

I would also like to thank my committee members, in particular Dr. Joshua Levine for his instruction in visualization foundations, which helped me build the background for completing this research. He was also available and willing to share his experiences when pursuing a PhD Degree. I would like to thank Dr. Brian Dean and Dr. Amy Apon for their suggestions concerning the techniques developed in this research.

In addition, I would like to thank my colleagues for their intelligent discussions about this research and their collaboration in conducting the cognitive studies. They are Dr. Mike Kirby, Dr. Ross Whitaker, Dr. Bill Thompson and Dr. Sarah Creem-Regehr, and graduate student Lace Padilla and Ian Ruginski from The University of Utah; Dr. Mary Hegarty and graduate student Alexander Boone from the University of Santa Barbara; Dr. Mahsa Mirzargar from the University of Miami; and Dr. Michael K. Lindell from the University of Washington.

Finally, I would like to thank Barbara Ramirez for her professional help in editing this dissertation.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 Motivation

The U.S. National Hurricane Center (NHC) begins issuing advisories every six hours at 04:00, 10:00, 16:00, and 22:00 U.S Eastern Standard Time when a tropical storm in either the Atlantic or Eastern Pacific region develops into a cyclone. These advisories take the form of several text documents, including the forecast advisory, which, along with other information, includes the predicted latitude and longitude of the storm center, the wind intensity, and the storm size. Additionally, an advisory may also contain a list of all current coastal watches and warnings, and possible storm tides.

To enhance the users' ability to understand these forecasts, the NHC also produces a suite of graphic products including the track forecast cone, the surface wind speed probabilities, and the surface wind field. Among these products, the track forecast cone, often referred to as the cone of uncertainty or uncertainty cone, is the most widely adopted by the news media, the general public, and the appropriate governmental officials. Figure 1.1 exemplifies the NHC cone of uncertainty for Hurricane Isaac 2012 [NOA12], with the black line indicating the forecast track of the center for the next three days. The orange and the black dots along this line represent the origination and the predicted succeeding centers. Capital letters label these dots to indicate the predicted intensities at various instances, with D, S, H, M representing tropical depression, tropical storm, hurricane, and major hurricane, respectively. The width of the cone covers a 67% confidence region through which the forecast track will pass. Hurricane watches and warnings are indicated by the pink and red areas along the coast line while tropical storm watches and warnings are indicated by the yellow and blue areas. The legends of these graphical elements and additional information such as time, advisory

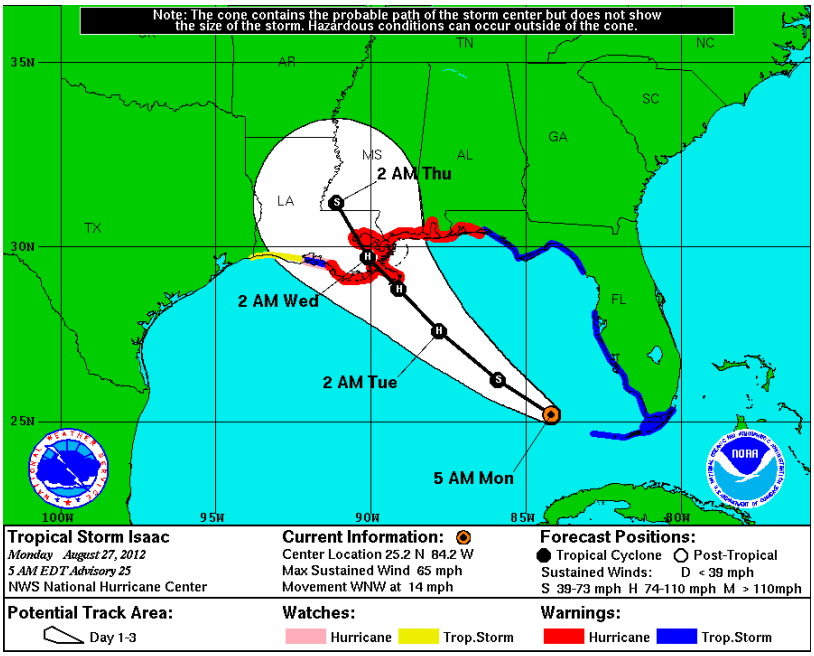



Figure 1.1: Example from a U.S National Hurricane Center track forecast cone: Hurricane Isaac at 5 AM EDT, August 25, 2012.

number, and wind speed are given in the box in the bottom of the image.

Even though the NHC cone of uncertainty is the most important current visualization of hurricane forecasts, much research conducted on its effectiveness has experimentally shown its intrinsic limitations. The initial study conducted by Broad et al. [BLWS07] analyzed the responses of the general public and the major media to Hurricane Charley 2004, reporting that this forecast cone and its underlying probabilistic concepts can be misinterpreted. For instance, an increase in the width of the cone tended to be easily mistaken as an increase in the size of the storm. In addition, people may focus on the black line indicating the possible track of the center of the hurricane mistakenly believing there is less danger further out. Finally, the track forecast cone is a binary inside-outside representation that can possibly give those outside the cone a false sense of security, while those inside tend to have an exaggerated feeling of danger. These conclusions were supported by Meyer et al. [MBOP13], who explored participants' responses to hurricane forecast information using Stormview Lab, a virtual environment that allows participants to search information on simulated hurricanes. Their results found that participants acted differently to a cone of uncertainty with or without the black line. Specifically, participants exhibited a higher level of preparedness when shown the former.

To address this emphasis, the NHC began providing a supplementary version of the cone of uncertainty without the black track line as seen in Figure 1.2. In addition, researchers outside of the NHC

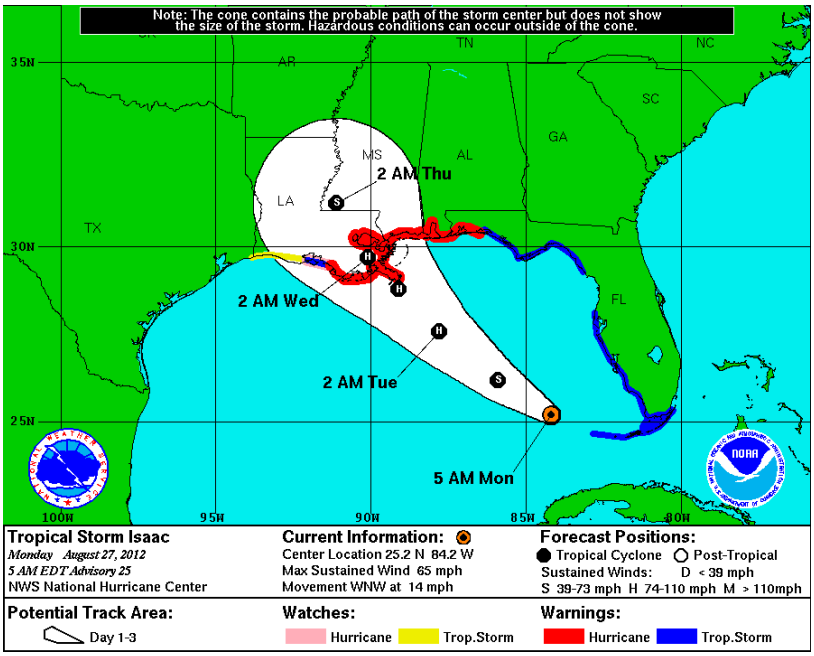



Figure 1.2: Example of a U.S. National Hurricane Center track forecast cone without track line: Hurricane Isaac at 5 AM EDT, August 27, 2012.

focused on developing better alternatives to this visualization. One important work was conducted by Cox et al. [CHL13], which implicitly depicts the hurricane-predicted spatial information and the uncertainty by producing and displaying an ensemble of hurricane-predicted tracks. To generate these tracks, their approach begins with the current location, speed, and bearing included in the NHC advisory. Using a Markov Random Field approach, they produced multiple potential tracks by predicting changes in these characteristics of a given hurricane using a combination of historical hurricane tracks dating back to 1945 in conjunction with the current NHC advisory. These tracks are continuously generated and drawn in such a way that the statistical distribution of the resulting ensemble closely matches that of the one implied by the cone of uncertainty. Because the tracks are overlaid on top of one another, the old tracks fade over time to keep the display from becoming cluttered; as the snapshot of this visualization in Figure 1.3 shows, individual tracks are drawn with a low opacity, which accumulates as tracks overlap one another, meaning areas with a higher concentration of tracks are more opaque than those with lower ones. Thus, more opacity indicates a higher level of certainty. Compared to the forecast cone, this technique not only preserves important statistical characteristics similar to those of the NHC hurricane advisory but also suggests the spatial uncertainty of a forecast by showing a wide range of predicted tracks both inside and outside the cone. To examine the effectiveness of this visualization, a subsequent user study was conducted, one in which participants were shown both this display

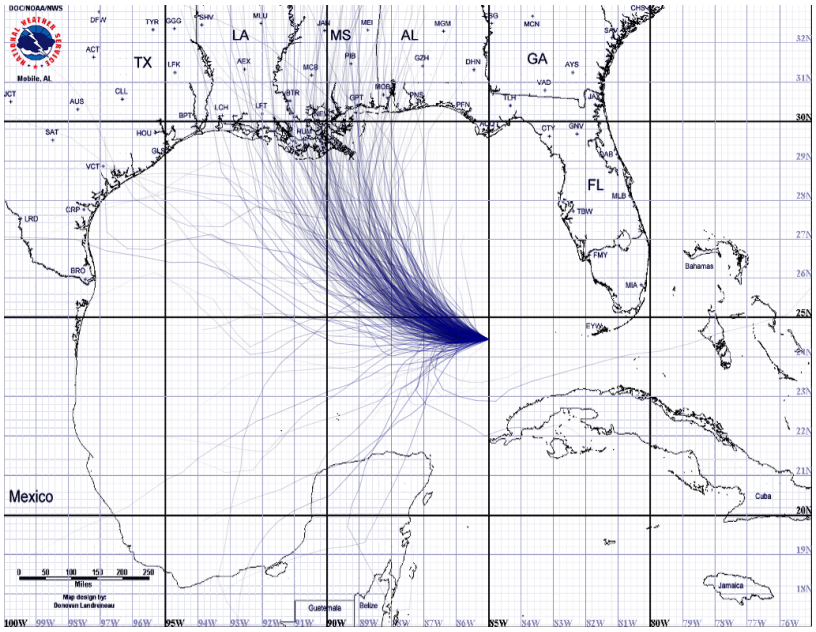



Figure 1.3: Example from direct ensemble display developed by Cox et al. [CHL13]: Hurricane Katrina at 10 AM CDT, August 27, 2012 (reprinted with permission from [CHL13]).

and the track forecast cone separately, and were then asked to estimate the strike probabilities of the sectors defined by the eight cardinal and ordinal directions. The results demonstrated that using this new method, subjects made similar predictions to the ones made when shown the cone, but with a clearer understand of the presence of outliers.

In addition, recent research conducted by Ruginski et al. [RBP+16] compared this ensemble display with the NHC track forecast cone and other summary displays derived from it as shown in Figure 1.4. They examined participants' estimations of potential damages to simulated oil rig platforms in the Gulf of Mexico at 24 and 48 hours into the prediction. The experimental results suggested that for predictions at 24 hours from the time of the advisory, little or no differences were found among the visualizations, while at 48 hours the ensemble display produced peak damage estimates significantly lower than the other summary displays. In addition, the curve of damage from the center of the prediction was shallower in the ensemble display than in the track forecast cone, reaching a much lower peak as seen in Figure 1.5, consistent with estimating a less likely but broader, spread of effects. A think-aloud session after the experiment in which participants were asked to describe their ranking strategies revealed all displays except the ensemble led to the misconception that the increasing size in the geometry meant a corresponding increase in the size the storm. A questionnaire administered after the trials showed that 69-81% of participants reported seeing the storm increasing in size

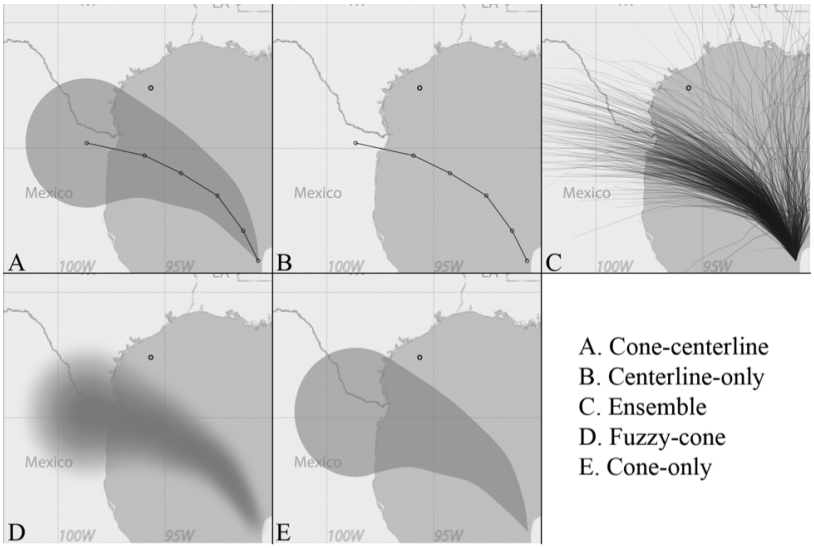
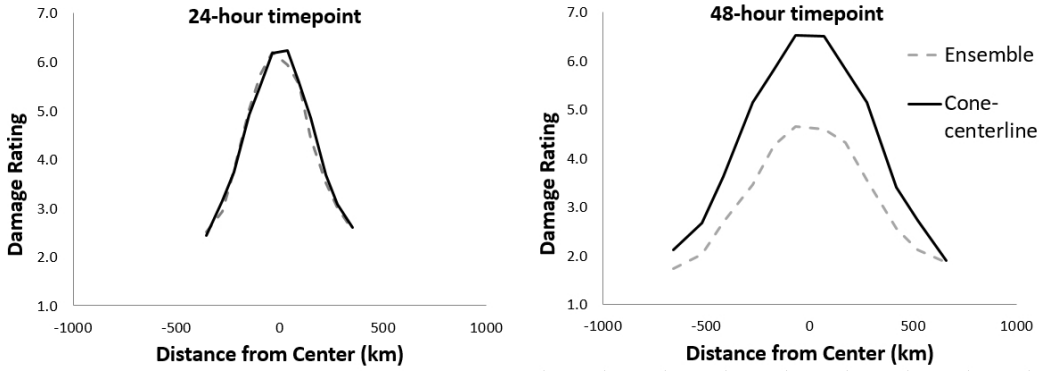



Figure 1.4: Ensemble visualization v.s. statical summary displays *Source*: [RBP+16]




(a) 24 hours

(b) 48 hours

Figure 1.5: Mean hurricane damage estimates as a function of distance from centerline, for ensemble visualization (dotted lines) and the NHC cone (solid lines) for the 24-hour and 48-hour time points. Data from the study conducted by Ruginski et al. [RBP+16].

and intensity in the cone-centerline, fuzzy-cone, and cone-only, but only 31% with the ensemble and 5% with the centerline only. These studies provide solid evidence that the direct ensemble visualization developed by Cox et al. [CHL13] is more effective than the NHC cone of uncertainty because it more accurately conveys the spatial distribution associated with prediction uncertainty, minimizing the confusion of the spatial attributes of a prediction.

However, this technique also has limitations. First, the predicted tracks were designed with no indication of time, meaning this visualization does not address important time and location-specific information such as the likelihood that the storm will hit an area with hurricane strength winds at a particular time on a particular date. Emergency managers responsible for advance planning for an oncoming hurricane are anxious to have such information readily available in visualizations (personal communication: Matthew Green, Federal Emergency Management Agency representative at the NHC, March 2014). Second, a direct ensemble visualization which displays all of the available data has the potential to become a confused jumble. This overdraw problem not only prevents the viewers from effectively understanding the information represented by the visualization but also prevents the designers from visually encoding supplementary storm characteristics using such conventional techniques for representing uncertainty as manipulating geometries and varying colors. For instance, polylines could be used with colors mapped to storm attributes, for example, the average intensity of the tracks. However, the variation of colors can be extremely difficult for the viewers to distinguish from a visualization with visual clutter.

The research reported here was motivated by these limitations of the direct ensemble display. Specifically, it focuses on developing ensemble-based visualizations to 1) support explorations of time- and location-specific information of hurricane predictions and the included uncertainty, and to 2) minimize visual clutter as much as possible.

## 1.2 Research Approach and Contribution

To realize these objectives, in this work we first developed a visualization encoding three levels of positional storm-strike risk for a specific point in time by generating and smoothly interpolating robust statistics from ensembles of hurricane forecast tracks developed by the NHC. Specifically, this approach begins by sampling each path from the ensemble at a specific time to create an ensemble of points fixed in time. The concept of simplicial depth was then applied to provide a centrality ordering of the time samples. Subsequently, this simplicial depth field was interpolated using an adaptive radial basis functions (RBF)

interpolation technique developed in this study. Finally, a geospatial visualization incorporating the concept of risk was designed based on the simplicial depth field as seen in Figure 1.6. Red, orange, and maroon were used to indicate the 33%, 66%, and 99% confidence intervals, respectively.

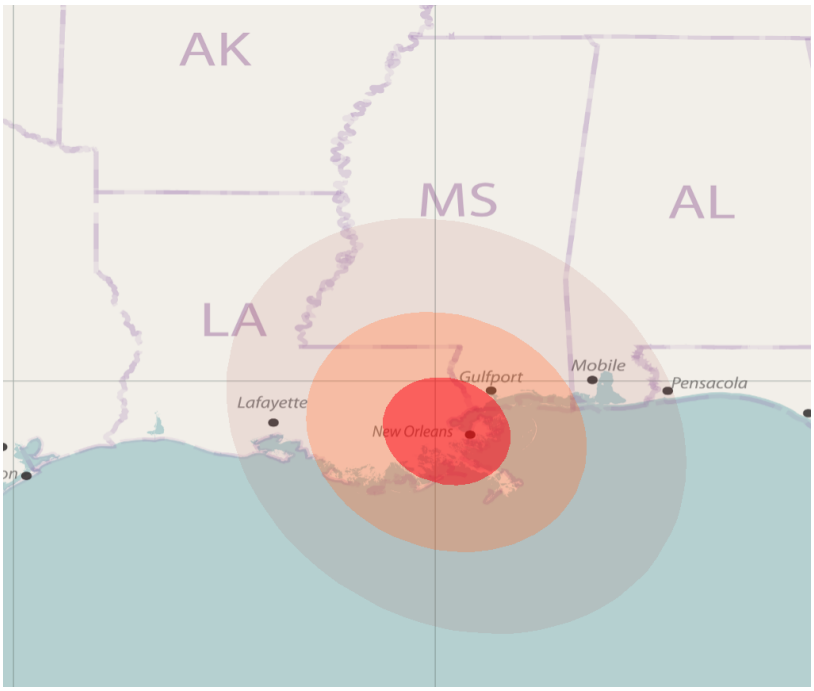
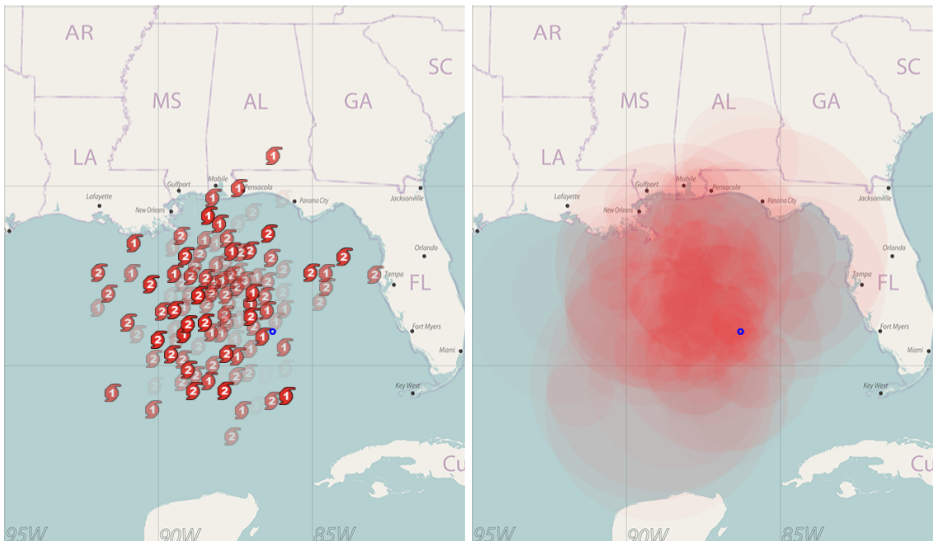



Figure 1.6: The visualization shows three levels of confidence intervals at 36 hours in the prediction: Hurricane Isaac at 5 AM EDT, August 27, 2012.

This visualization was shown to the meteorologists at the NHC and other experts in the visualization community; one important resulting critique was that as the confidence intervals, which were summary displays, grew over time, it was easy to misinterpret the increase in the intervals as an increase in the storm size. To address this issue, we developed an alternative visualization explicitly displaying a subset of samples taken from the ensembles of the hurricane-predicted locations. The members of this subset accurately preserved the original spatial distribution and maintained appropriately defined spatial separations. Similar to the confidence interval visualization, it began with extracting the ensemble of points fixed in time from the NHC paths ensemble. These points were then projected to a space where they were uniformly distributed, referred to as the UD space. Subsequently, a subset of the samples was selected to achieve a Poisson Disk distribution in this UD space. Finally, the selected points were transformed back into the original space and displayed geospatially. To demonstrate that this approach can be used to annotate geospatial displays with additional hurricane information, this study produced two animated visualizations. One of these conveys the

hurricane intensity using glyphs designed by the NHC at the sample points (Figure 1.7a), while the other implies storm size using circles, the radii of which are determined by the storm size (Figure 1.7b). To verify




(a) glyphs (b) sizes

Figure 1.7: Two animated visualizations generated by representatively sampling the hurricane prediction ensemble.

their effectiveness, a subsequent cognitive study was conducted in collaboration with researchers from the University of Utah and the University of California, Santa Barbara comparing these visualizations to summary displays by examining participants' estimations of potential storm damage. The experiment results reported that by shown the ensemble display using glyphs, participants were less likely to associate the storm damage to the distance from the center of the forecasts than shown the other visualizations, suggesting that this visualization is interpreted more as a uncertainty distribution rather than being incorrectly comprehended that the hurricane would be strongest and largest at the center of the forecast.

Even though this ensemble display is an improvement of implying hurricane forecast uncertainty, it also has limitations. Most importantly, this approach requires generating separate displays for representing individual storm characteristics at multiple points in time, e.g. visualizing predicted storm size and intensity over 120 hours requires 240 animated visualizations. This large number of displays places heavy cognitive load on viewers, making it difficult for them to comprehend the information precisely. In addition, previous studies have reported that using separate displays creates more perceptually and cognitively work for viewers [Har03], and significantly prevents them from discovering and analyzing patterns of multiple variables [HTER04]. This research addresses these issues by generalizing representative sampling to process

ensembles of forecast tracks, developing a visualization by directly rendering a small number of representative paths extracted from the original ensemble. To compute these representative paths, this approach extracts the spatial median path from an ensemble, equally partitioning it into two groups, then constructing more paths by recursively finding medians from succeeding groups. Since individual paths significantly represent their corresponding subdivision of the original ensemble, they can accurately depict its spatial distribution at a certainty level of confidence. Further, information of other storm characteristics are assigned to these paths using the RBF interpolation technique developed previously. Since the number of tracks needing to be rendered in the display has been reduced, they can be used for portraying the forecast without leading to visual clutter. Figure 1.8 exemplifies a visualization built based on this technique, where hurricane tracks are indicated using polylines the colors of which are associated to time-specific storm intensities, and circles indicating storm sizes are placed on predicted locations along these tracks. Subsequently, a cognitive study

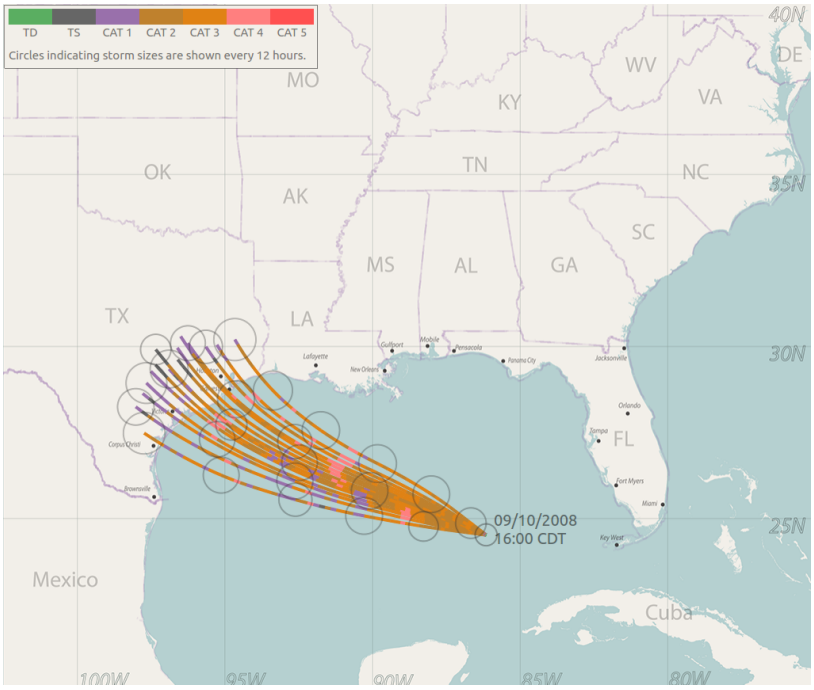



Figure 1.8: The visualization portrays prediction of Hurricane Ike at September 10th 2008 by showing 15 representative forecast tracks the colors of which are associated to predicted storm intensities. The predicted storm sizes are indicated by circles placed along these tracks.

examining viewers' interpretation of this visualization has been conducted. The preliminary experimental results revealed that this visualization was interpreted by the participants more as an uncertainty distribution rather than misunderstanding that the hurricane was predicted to be strongest in the spatial center of the forecast, thus effectively enhancing viewers' estimation of potential risks.

## 1.3 Thesis Organization

Chapter 2 discusses the theories, concepts and techniques related to the work reported here, including the NHC forecast process and its underlying uncertainty, emphasizing those specifically for hurricane forecasts and forecast ensemble. Chapter 3 explains the RBF-based approach developed for interpolating storm characteristics as well as a visualization conveying possible hurricane strikes using multiple levels of confidence intervals. Chapter 4 introduces an algorithmic framework for computing a representative subset from an ensemble of points, showing how this approach can be used to augment visualizations of time- and locations-specific hurricane predictions. Chapter 5 extends representative sampling to process more general ensembles of forecast tracks, developing a hurricane visualization encoding both the spatial information and information of other storm characteristics over a forecast period of time. Chapter 6 concludes this research by summarizing its impacts to the filed of hurricane forecasting, and discussing its potential benefits to the general visualization community.

# Chapter 2

# Background

## 2.1 NHC Forecast Process and the Underlying Uncertainty

The ultimate goal of a hurricane forecast is to predict the storm's behavior as precisely as possible including its path, size, and intensity, thus guiding effective decision-making for hurricane response to minimize the loss of life and property [Ste09]. However, according to reports from the NHC [Lou14, NOA15a], uncertainty is inevitable during the entire forecasting process. The NHC begins forecasts by collecting the current atmospheric status from all available observations, including satellites, reconnaissance aircrafts, ships, buoys, radars, automated surface observation systems, and coastal monitoring systems. These data collected contain errors because measurements and their results are inherently imprecise [NIS92]. In addition, the measurement value of an attribute, for example air pressure, is highly dependent on the approaches or methods used. For this reason, an estimation, for example the mean of the various values obtained from the instruments such as electronic, mercury, or aneroid barometers, is used. Further, the space and time resolutions of observations are limited by the number of devices employed, meaning thus again approximations are used. For example, only a certain number of sensors measuring wind speed can be deployed over a region so that the data at locations without a sensor are evaluated by interpolating results from the measurements taken in their vicinity.

The situation is further complicated by the uncertainty introduced when using the data first collected to initialize a set of forecast models, usually referred to as numerical weather predicted (NWP) models. For example, a wide spectrum of atmospheric physics and chemistry, such as atmospheric electricity, clouds, small-scale dynamics, and convection, are understood incompletely by scientists, meaning that these models

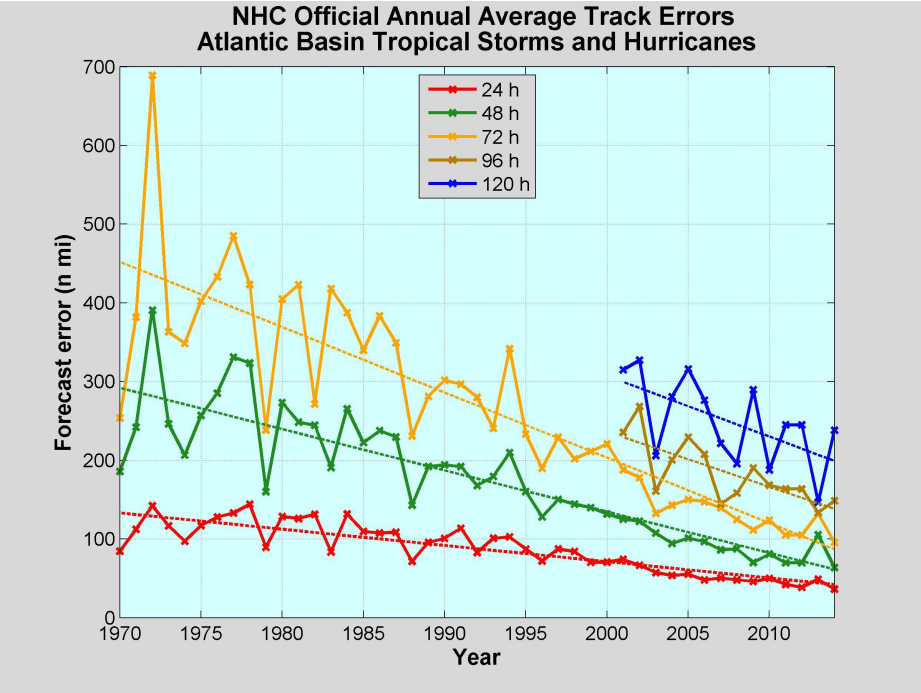



Figure 2.1: NHC official annual average track errors for Atlantic Basin tropical storms and hurricanes, *Source*: [NOA15b].

are imperfect, thus producing results with less than 100% accuracy. These NWP models accumulate error from data collected, subsequently evolving to predict the future development of the storm, this integration further increasing the uncertainty. This situation can be seen in Figure. 2.1 which shows NHC official annual average track errors for Atlantic basin tropical storms and hurricanes published in 2015, with the horizontal axis indicating years while the vertical axis indicates forecast errors measured in nautical miles. Errors at 24, 48, 72, 96, and 120 hours from the beginning of a forecast, color-coded in red, green, orange, brown, and blue, respectively, indicate less accuracy as forecast time increases; for instance, the error at 24 hours in 2010 was less than 100 nmi while that at 120 hours in the same year increased to approximate 250 nmi.

Simulation results from multiple NWP models are the basic components used by meteorologists for producing the final forecast. However, individual models vary in structures and complexity, possibly leading to diverse forecasts. For instance, one type of taxonomy of these models categorizes them as dynamical and statistical. Dynamical models that solve the physical equations governing atmospheric motions simulated by using high performance computers are sophisticated and more accurate, while statistical models, developed based on an understanding of the relationship among storm characteristics and behaviors, are cheaper to compute and include more uncertainty. Details of the advantages and limitations of the NHC forecast models

available are discussed in a technical report on its website [NOA09]. To address the uncertainty of using various models, ensemble approaches, also referred to as consensus models, are usually used by meteorologists [Lew05, SDR00, Atg99, MWK14, WMK13, Mur88, MBPP96, TK93, TK97]. These methods include both multi-model ensembles obtained by running multiple NWP models with the same initial state to estimate the atmospheric evolution and perturbed parameter ensembles generated by running a single model multiple times, each simulation run beginning with slightly perturbed initial conditions that are combined to produce a forecast. The simplest way to form a consensus forecast is to take all possible members of an ensemble and compute the average. More sophisticated corrected-consensus models, assigning different weights to each member, are also used by the NHC. While results from consensus models are generally considered to be more accurate than results from individual models, they involve uncertainty accumulated from the imprecise initial observations and the intrinsic defects of NWP models and, thus, cannot be directly issued as an official forecast. Therefore, forecasters need to evaluate predictions from these models by comparing them to the current atmospheric status and their experiences, producing a forecast that introduce another degree of uncertainty. These uncertainties are summarized as tables of the NHC official annual average track and intensity forecast errors [NOA17b, NOA17a]. Individual items of these tables describe the track/intensity errors at 0, 12, 24, 36, 48, 72, 96, 120 hours in the prediction at specific years. For stable purposes, the NHC uses a five-year sample to determine the current forecast errors [NOA17c]. Typically, the final forecast and its underlying uncertainty are issued as a hurricane forecast advisory, as shown in Figure. 2.2, where the errors of this forecast are highlighted by red in this example.

```
TROPICAL STORM ISAAC FORECAST/ADVISORY NUMBER 24
NWS NATIONAL HURRICANE CENTER MIAMI FL AL092012
0300 UTC MON AUG 27 2012

**********

A HURRICANE WARNING IS IN EFFECT FOR...
* EAST OF MORGAN CITY LOUISIANA TO DESTIN FLORIDA...INCLUDING
METROPOLITAN NEW ORLEANS...LAKE PONTCHARTRAIN...AND LAKE MAUREPAS

**********

TROPICAL STORM CENTER LOCATED NEAR 24.2N 82.9W AT 27/0300Z
POSITION ACCURATE WITHIN 25 NM

PRESENT MOVEMENT TOWARD THE WEST-NORTHWEST OR 285 DEGREES AT 12 KT

ESTIMATED MINIMUM CENTRAL PRESSURE 993 MB
MAX SUSTAINED WINDS 55 KT WITH GUSTS TO 65 KT.
50 KT.......  60NE 0SE 0SW 60NW.
34 KT.......180NE 120SE 80SW 180NW.
12 FT SEAS..200NE 200SE 150SW 150NW.
WINDS AND SEAS VARY GREATLY IN EACH QUADRANT.

FORECAST VALID 27/1200Z 25.2N 84.8W
MAX WIND 60 KT...GUSTS 75 KT.
50 KT...  60NE 0SE 0SW 60NW.
34 KT...180NE 120SE 80SW 180NW.

FORECAST VALID 28/0000Z 26.5N 86.7W
MAX WIND 70 KT...GUSTS 85 KT.
64 KT...  20NE 20SE 0SW 20NW.
50 KT...  60NE 60SE 30SW 60NW.
34 KT...180NE 120SE 80SW 180NW.

FORECAST VALID 28/1200Z 27.9N 88.3W
MAX WIND 80 KT...GUSTS 100 KT.
64 KT...  30NE 20SE 10SW 20NW.
50 KT...  60NE 60SE 40SW 50NW.
34 KT...180NE 150SE 80SW 160NW.

FORECAST VALID 29/0000Z 28.9N 89.3W
MAX WIND 85 KT...GUSTS 105 KT.
50 KT...  70NE 60SE 50SW 60NW.
34 KT...180NE 150SE 90SW 160NW.

FORECAST VALID 30/0000Z 30.6N 90.3W...INLAND
MAX WIND 70 KT...GUSTS 85 KT.
50 KT...  50NE 60SE 60SW 40NW.
34 KT...100NE 150SE 100SW 90NW.

EXTENDED OUTLOOK. NOTE...ERRORS FOR TRACK HAVE AVERAGED NEAR 175
NM ON DAY 4 AND 225 NM ON DAY 5...AND FOR INTENSITY NEAR 20 KT EACH
DAY

OUTLOOK VALID 31/0000Z 33.0N 91.0W...INLAND
MAX WIND 30 KT...GUSTS 40 KT.

OUTLOOK VALID 01/0000Z 36.0N 91.0W...INLAND
MAX WIND 20 KT...GUSTS 30 KT.
```

Figure 2.2: Example from a US NHC Advisory: Hurricane Isaac at 0300 UTC, August 27, 2012.

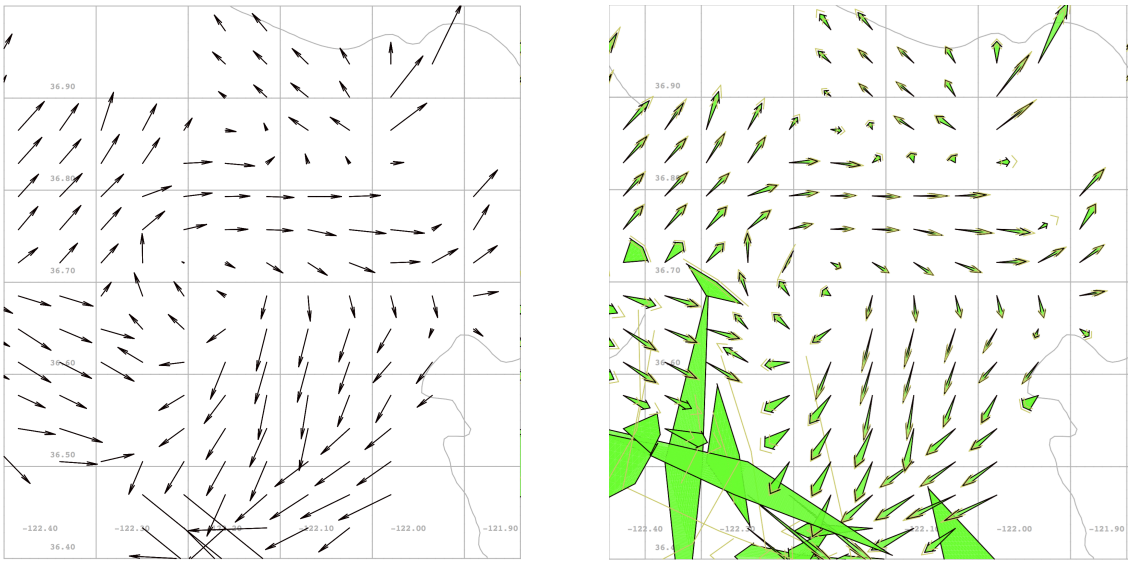
## 2.2 Uncertainty Visualization

Understanding uncertainty in hurricane forecasts is important because it plays a significant role in enhancing hurricane emergency decision-making. For instance, when a hurricane approaches an area, the local emergency managers receive forecasts with an additional statement of their underlying uncertainty from the NHC, then determine whether to issue an evacuation alert to the public by considering both the evacuation cost and the possible hurricane damages. In general, these managers have to prepare for a potential serious situation, but they do not want to over-prepare, spending money unnecessarily. Techniques for improving interpreting hurricane forecasts by including this uncertainty can address this issue. Specifically, uncertainty visualization has the potential to provide a more complete and accurate explanation of data for the viewers to analyze, functioning as an asset to this field.

In general, the visualization community categorizes uncertainty due to its source [PWL96, Pan08, Pot10]. The first taxonomy in this area, which is also the one most cited, was introduced by Pang [PWL96], who categorized uncertainty into acquisition, transformation, and visualization. Acquisition uncertainty occurs during data collection through observations, measurements, and numerical simulations, while the transformation is introduced by applying transformational operations such as rescaling, re-sampling, and quantizing on these raw data, and the visualization is generated during the process of visualizing the data. Specifically, in connection with the hurricane forecast process, errors result from simulating the complex atmosphere since approximations are frequently used when measuring the environmental status, i.e. the wind speed, the air pressure, and the air temperature. These approximations and errors are acquisition uncertainty. The data collected need to be post-processed, such as converting from one unit to another, filling in missing data through interpolation, extracting a continuous field from the raw data and re-sampling to support the utilization of NWP models, thus resulting in transformation uncertainty. In addition, to determine the final forecast, meteorologists turn multiple simulation results into visual representations, such as images and animations, where the varied rendering attributes of colors, opacities, and geometries affect the interpretation of the forecast. For example, rendering data using a bright red can leave a deep impression on the viewers, potentially over-emphasizing the information encoded.

In addition to classifying uncertainty, Pang [PWL96] also compiled a list of fundamental approaches for presenting uncertainty, including adding glyphs and geometry, modifying rendering attributes, and using animation, sonification, psycho-visual. As the last two are beyond the scope of this research, we review only the first four. A glyph is a visualization unit that encodes information in its shape and color. Figure 2.3

exemplifies an application of using glyphs to enhance viewers' understanding of the vector field of ocean currents. Conventionally, arrows are used to represent the bearing and magnitude of these currents as seen in



(a) (b)

Figure 2.3: Visualizing ocean currents using glyphs: (a) uses arrow glyphs and (b) uses uncertainty glyphs. *Source*: [WPL96].

Figure 2.3a. However, this type of glyph indicates nothing about the uncertainty of these two parameters. To address this limitation, Wittenbrink et al. [WPL96] developed a more sophisticated glyph that encodes these additional uncertainties as seen in Figure 2.4. Specifically, the width of the arrow heads depicts the range of

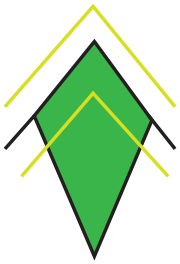


Figure 2.4: Example of the uncertainty glyph designed by Wittenbrink et al. [WPL96].

possible bearings at that location, while the additional arrow heads, which are color-coded yellow in Figure 2.4, indicate the range of possible magnitudes. Figure 2.3b visualizes the same vector field represented in Figure 2.3a. The latter method has been experimentally shown to be superior to the former since it portrays

more information that can be accurately interpreted by the viewers [WPL96].

While glyphs are used by many applications, they can only represent discrete data. To portray continuous data, designers usually display uncertainty by adding geometry such as contour lines, isosurfaces, and streamlines. In general, this approach is combined by modifying the rendering attributes of geometry, meaning that varying shading and pseudo-colored parameters are used to indicate levels of uncertainty. For example, Kai et al. [PH11] used isolines to represent a temperature field ensemble of 25° C and its included spatial uncertainty. As seen in Figure 2.5, the isoline of statistical mean is drawn in solid black while the areas experiencing that particular temperature with different probabilities are color-coded using the color scheme shown by the legend on the right of the image. Figure 2.6 exemplifies the technique developed by Fersti

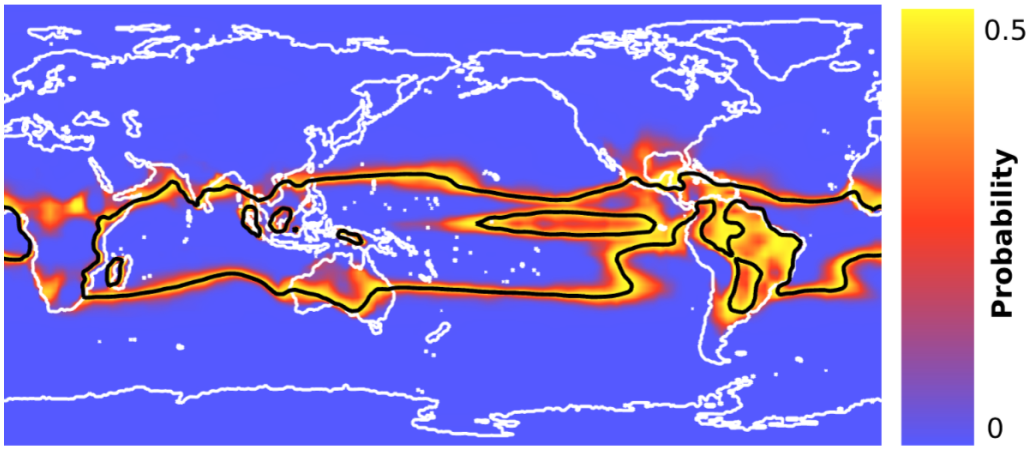
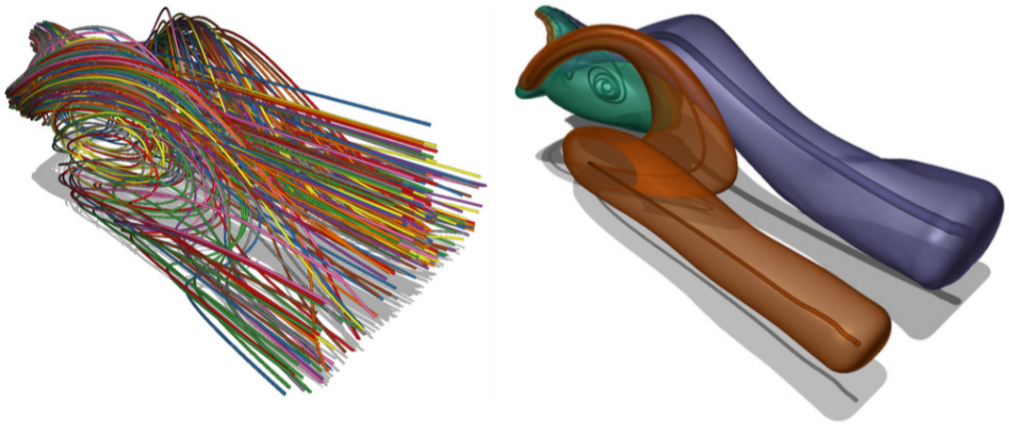



Figure 2.5: Example of using isolines to represent the spatial uncertainty of a temperature field ensemble of 25° C, *Source*: [PH11].

et al.[FBW16], referred to as streamline variability plots which represent the spatial uncertainty of a vector field of 3D flows using contours. Specifically, this technique begins by computing the principle component analysis (PCA) space of a collection of streamlines extracted from the vector field (Figure 2.6a) and clustering them in the PCA space. Subsequently, each of these clusters is represented by a track of the statistical mean and a contour indicating a 50% confidential interval as seen in Figure 2.6b. A color scheme is implemented to distinguish the various clusters.

In addition to these approaches depicting uncertainty using statical images, animation is also an effective method because of its unique advantage of mapping information to such visual attributes as speed or duration, motion blur, and range. For instance, Brown [Bro04] represented the spatial uncertainty of terrain data using oscillation techniques, meaning the error describing the possible range of the height at each vertex of the data was portrayed by smoothly varying the Z value of the vertex in the display as a function of time.



(a) 3D flow streamlines (b) variability plot

Figure 2.6: Visualizing 3D flow streamlines using the variability plot developed by Ferstl et al. [FBW16]. *Source*: [FBW16].

Figure 2.7 exemplifies two frames of an animation using this technique where the representation of spatial uncertainty is highlighted within the dashed rectangle. This technique illustrates the blur metaphor which

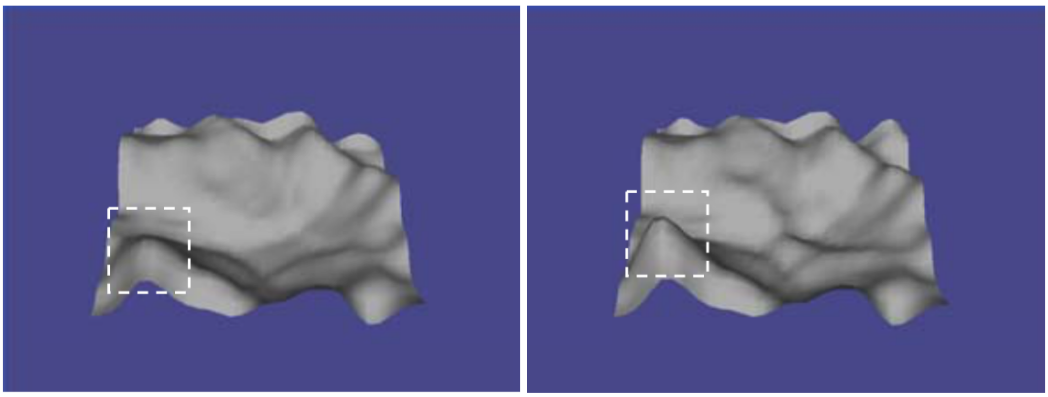


Figure 2.7: Frames of an animation of the oscillation technique developed by Brown [Bro04]. *Source*: [Bro04].

directly presents the accuracy of the data, requiring less cognition from the viewer to understand.

## 2.3 Uncertainty Visualization of Hurricane Forecast

The initial visualizations of hurricane forecasts, developed in 1999 as the result of a collaboration among the Climate Prediction Center, the NHC, and the Hurricane Research Division, displayed only track information for an entire season without any depiction of its uncertainty. This visualization super-imposed

color-coded polylines indicating tropical depressions, tropical storms, and hurricanes on satellite images. Figure 2.8 shows such a display for the Atlantic Hurricane Season of 1999 with these storms being represented by green, yellow, and red curves. As these images can easily be distributed, they were quickly adopted by

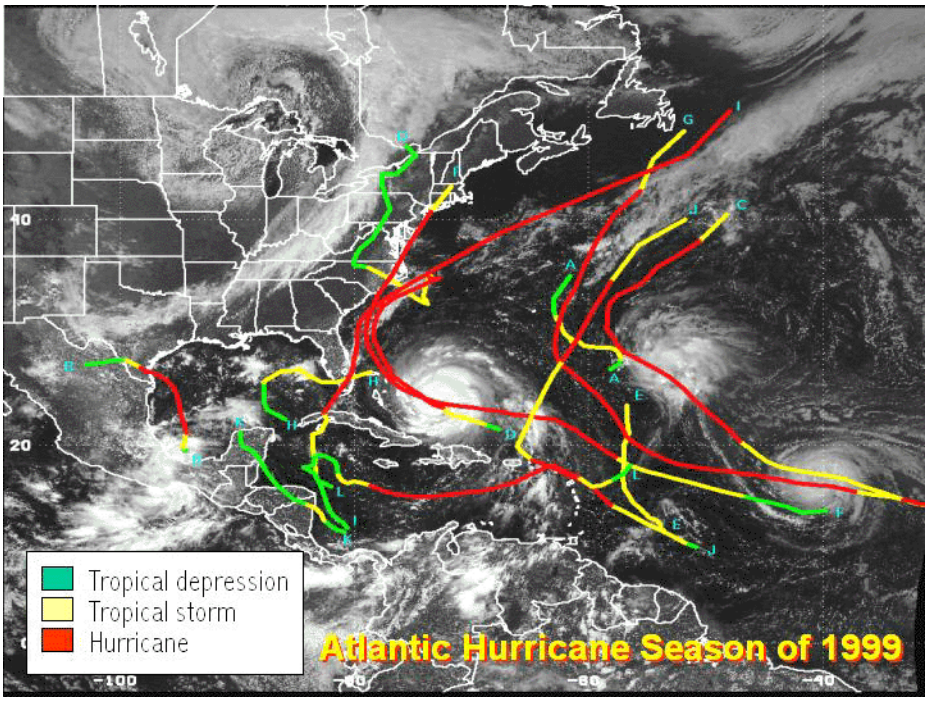



Figure 2.8: The earliest visualization of the Atlantic Hurricane Season of 1999 where color-coded polylines indicates tropical depressions, tropical storms, and hurricanes. *Source*: [NOA99].

the news media and broadcast to the general public. Subsequent research indicated that these images helped reduce U.S. deaths from hurricanes [BLWS07]. While this type of image was commended by both the general public and the appropriate governmental officials, it was quickly replaced by improved visualizations because of its inability to depict the uncertainty of the forecast tracks and to represent graphically in a single image increasingly complex data in the forecast.

To address these limitations, two visualizations have been developed, one by the NHC and one by the European Center for Medium-Range Weather Forecasting (ECMWF). The NHC developed the Surface Wind Field graphic to depict the areas potentially affected by the sustained winds of a storm and the coastal areas under storm watches and warnings. As seen in Figure 2.9, a dark red circle indicates a region may experience winds of hurricane force (74 mph) while an orange circle indicates a region with winds of tropical storm force (39 mph). Areas under hurricane and tropical storm warnings and watches are color-coded using light red, yellow, pink, and light blue, with the black dashed line indicating the track of storm to date and the

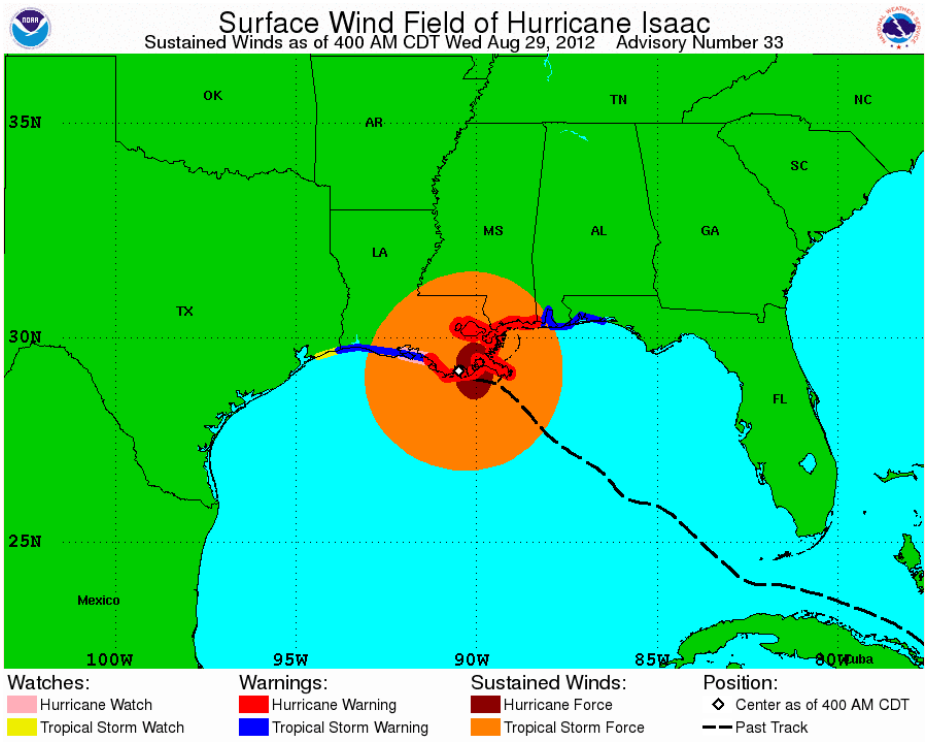



Figure 2.9: The Surface Wind Field visualization of Hurricane Isaac 2012 *Source*: [NOA00].

white dot the current location of the center of the storm. Additional information such advisory number and time are given at the top of the display. This visualization uses color and geometry to imply uncertainty. For instance, a warning indicates a higher level of confidence than a watch; thus, an area experiencing a Category 1 hurricane with a high probability is represented by the bright red band in the display.

In addition, scientists from the ECMWF developed a visualization of the strike probability of a forecast, defined as the probability of a forecast track passing within a 120 km radius of an instance during the next 120 hours. As seen in Figure 2.10, color-coded circles and polygons are used to indicate areas under different strike probabilities of a forecast. Since the strike probability is based on an ensemble forecast approach containing the more detailed spatial information included in the forecast, the NHC posts this visualization as well as the Surface Wind Field graphic on its website for consideration.

Even though these two visualizations integrated additional data from forecasts, they are separate displays, requiring the viewers to toggle between them to obtain information. In addition, forecasting techniques were becoming increasing more accurate and sophisticated, with a larger volume of data available for representations. Thus, the NHC further improved its visualization technique, developing the well-known track forecast cone introduced in Section 1.1. To construct the cone, a set of circles are placed along the

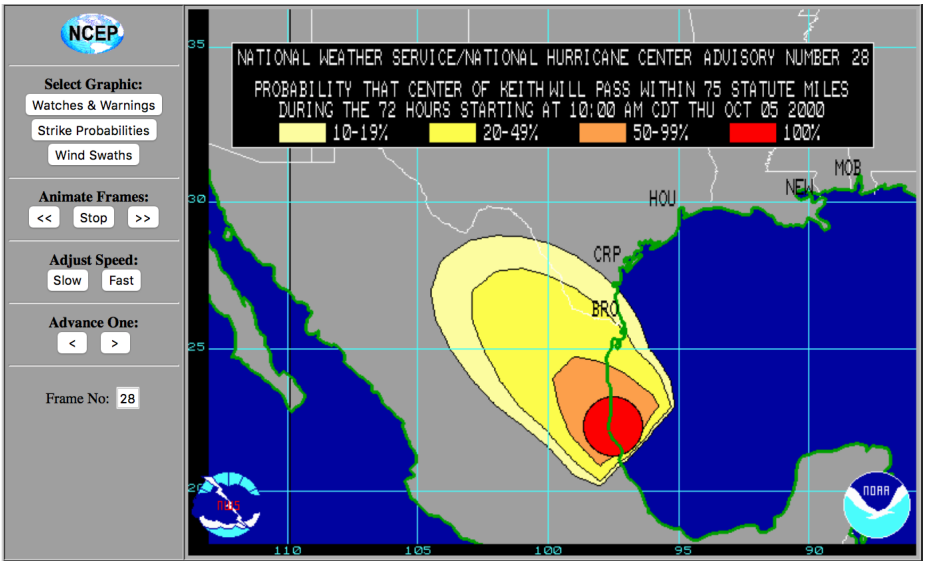



Figure 2.10: The visualization of strike probabilities of Hurricane Keith 2000. *Source*: [NOA00].

forecast track of the center of the storm (at 12, 24, 36 hours, etc.), the radii of which are specified, based on historical official forecast error over a 5-years sample so that the actual storm position will have a 67% possibility of falling within it. The cone is formed by enclosing the area swept out by these circles and, thus, implies the 67% confidential region through which the track will pass. The transparent grey area in Figure 2.11 exemplifies such a circle at 48 hours of Hurricane Ivan 2004.

Even though the track forecast cone has been widely adopted by the general publics, the news media, and the appropriate governmental officials, its statistical concepts tend to be easily misinterpreted by the viewers. More importantly, the storms vary in intensity and size in such a way that their effects can be experienced away from their centers. However, the track forecast cone does not consider these two types of uncertainties. To address this limitation, the Center began to provide additional text files describing the ensembles of predictions with additional time- and location-specific probabilities of surface (10-meter elevation) sustained (1 minute on average) winds for three speed thresholds (34, 50, and 64 kt) [NOA14]. Two types of probabilities are included in this text product. The first of them is the cumulative probability implying the overall chance that the storm wind considered will be experienced at any location over the regional map during such time periods as 0-12 hours, 0-24 hours, 0-36 hours, 0-48 hours, 0-72 hours, 0-96 hours, and 0-120 hours, while the second is the onset probability indicating the chance that the wind considered will be experienced at any location over the regional map during each of the following periods: 0-12 hours, 12-24hours, 24-36 hours, 36-48 hours, 48-72 hours, 72-96 hours, and 96-120 hours. To enhance viewers' ability to understand these probabil-

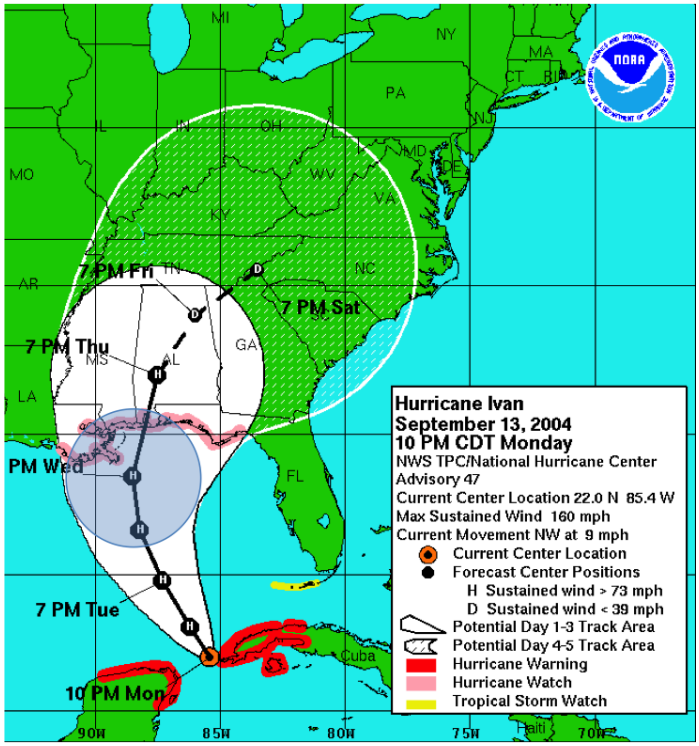



Figure 2.11: Example of a NHC track forecast cone: Hurricane Ivan at 10 PM CDT September 13, 2014.

ities, the NHC developed graphics for them, referred to as Wind Speed Probabilities. Figure 2.12 exemplifies such a visualization of the 0-120 hour interval of tropical storm strength (34 kt) winds for Hurricane Issac 2012 based on Advisory 27 issued at 1:00 PM CDT August 27. The white dot indicates the current location of the center of the storm, while the variations of probability are color-coded as demonstrated by the legend at the bottom of the image. Compared to the track forecast cone, this visualization provides the time- and location-specific information of winds of such familiar thresholds as tropical storm force and hurricane force and, thus, has more direct meaning on and impact on viewers. However, this approach inevitably requires producing multiple separate images to include wind at a specific speed during a specific period of time. It has been experimentally proven that using separate plots can significantly undermine viewers' ability to discover and analyze the patterns of the multiple variables conveyed by the entire data set [HTER04]. Additionally, since these probabilities are spatially sampled on a grid, they are coarse-grained and thus, subject to artifacts resulting in under-sampling.

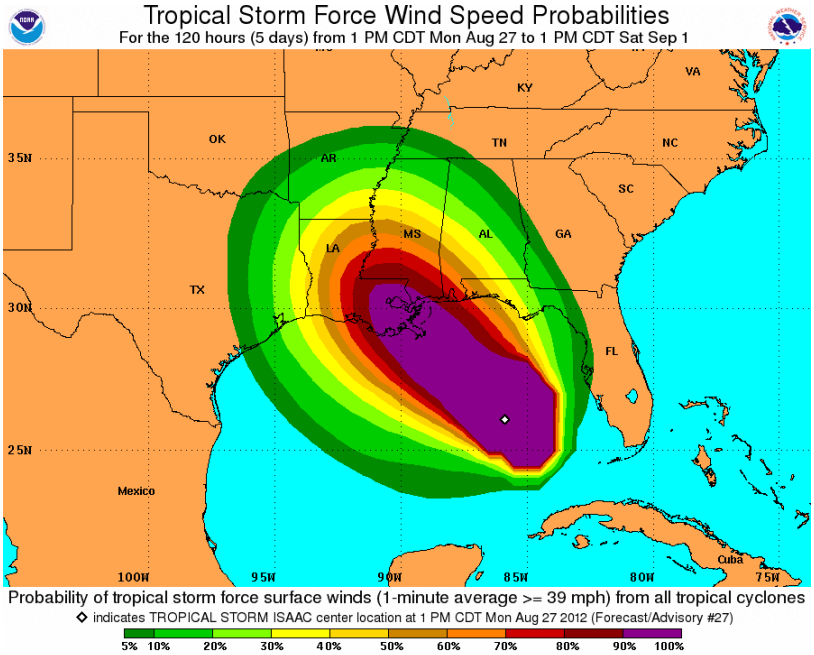



Figure 2.12: Example from a NHC Tropical Storm Force Wind Speed Probability: Hurricane Issac at 1 PM CDT August 27, 2012.

## 2.4 Uncertainty Visualization of Forecast Ensemble

In the context of hurricane forecasting as well as in other fields of weather predictions, ensembles have been widely utilized by meteorologists to simulate complex systems and to learn about the uncertainty included [Lew05, SDR00, Atg99, MWK14, WMK13, Mur88, MBPP96, TK93, TK97]. For instance, meteorologists forecast atmospheric conditions by designing numerical prediction models, typically nonlinear dynamical ones, the evolution of which are sensitive to the initial conditions. However, these conditions are usually determined by estimations because of the intrinsic inaccuracy of instruments and the incomplete understanding of the complexity of the atmosphere. Monte Carlo approaches are one of the most effective types of techniques that can represent these uncertainties by appropriately computing the probability density function (PDF) of nonlinear systems. These approaches first sample the PDF at the initial time and then derive the sampled initial state using prediction models to estimate the evolution of the system, while individual models are run one or multiple times by perturbing the initial conditions. The outcomes are commonly referred to as an ensemble and the individual elements as the ensemble members. The primary type of ensembles adopted by the research reported here is generated by the NHC for supporting the construction of the Wind Speed Probability graphics. These ensembles were generated using a Monte Carlo model and each

of these ensemble contains 1,000 predicted hurricane tracks over the 120 hours [DJK$^+$09, DKB$^+$13]. More specifically, this Monte Carlo model begins with the official track and intensity forecast of a hurricane and then, randomly sampling 1,000 errors from the previous five-years official track and intensity errors, which are formatted as tables and publicly available on the NHC websites[NOA17b, NOA17a]. Finally, these errors are added to the official forecast to produce 1,000 time-sampled hurricane tracks. Given the predicted locations and intensities, this model estimates the wind structure, which is defined as the radii of the 34-, 50-, and 64-kt winds in the four ordinal directions, i.e. NE, SE, SW, and NW, along each of these tracks using the statistical climatology and persistence (CLIPER) model [KSD$^+$07]. Typically, the generated ensemble is issued as a table, every 120 consecutive items of which represent a single forecast track, and individual items describe the hurricane-predicted locations, intensities, and radii of the three types of winds along the track at individual points in time. Figure 2.13 exemplifies a visual depiction of such an ensemble, where the thick polyline represents official track forecast of the hurricane; the thin polylines represent the 1,000 generated hurricane tracks; and hurricane-predicted intensities are color-coded.

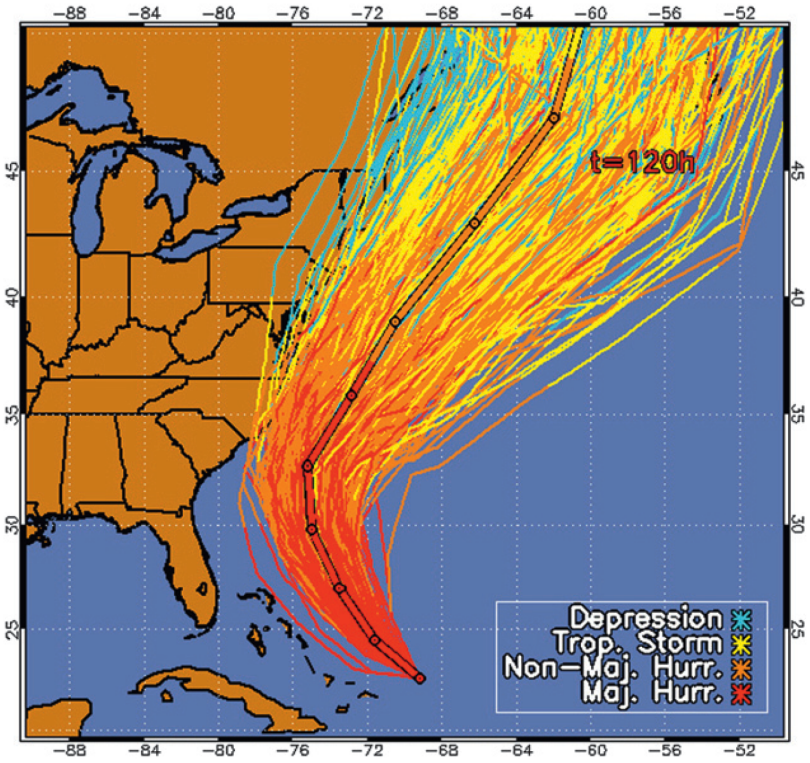



Figure 2.13: Example of an ensemble of 1,000 hurricane-predicted tracks generated by the NHC: Hurricane Earl, 0000 UTC, September 1, 2010. *Source*:[DKB$^+$13]

Because of the efficiency of using ensembles to understand hurricane forecasts and the underlying

uncertainty, many studies focusing on developing novel visualizations of them have been conducted. The initial one developed by Cox et al. [CHL13] was fully discussed in Section 1.1. In addition, Whitaker et al. [WMK13] developed contour band depth, enabling statistical analysis of isocontours extracted from scalar fields, subsequently using this to estimate median, order statistics, and outliers for drawing what they call contour boxplots. Based on this techqniue, Mirzargar et al. [MWK14] derived statistical characteristics from ensembles of multivariate curves extracted from flow fields, allowing them to draw curve boxplots. As one potential application, they demonstrated how ensembles of hurricane forecast tracks can be summarized using their method. Figure 2.14 exemplifies such a visualization of the ensemble for Hurricane Katrina 2005 generated using Cox et al.'s method. The yellow curve represents the track of the statistical median while the

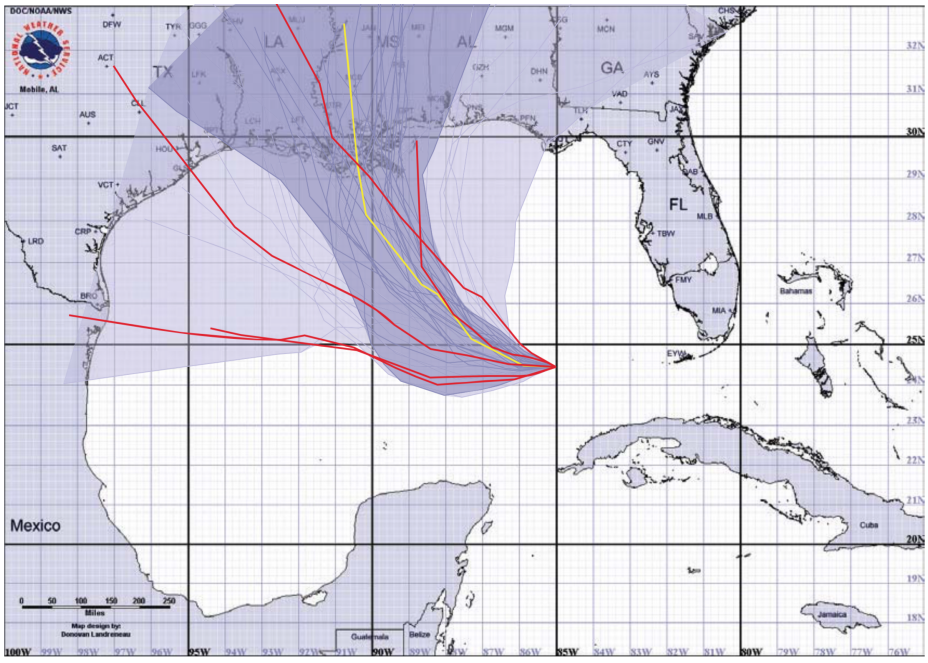



Figure 2.14: Example of visulization hurricane prediction ensemble by using curve boxplot developed by Mirzargar et al. [MWK14]: Hurricane Katrina 2005.

red curves indicate the outliers in their analysis. The dark blue region analogizes the 25% to 75% interval in a conventional boxplot, while the light blue region shows the area that statistically contains all tracks except the outliers. More recently, Ferstl et al. [FBW16] developed the streamline variability plot technique. Their technique applied a cluster analysis to hurricane predicted tracks abstracted as points in a PCA space. Subsequently, a Gaussian fit was utilized to compute a 50% confidence interval for each cluster. As seen in Figure 2.15, individual clusters are visualized as a combination of a track of statistical median and a colored region enclosing the confidence interval.

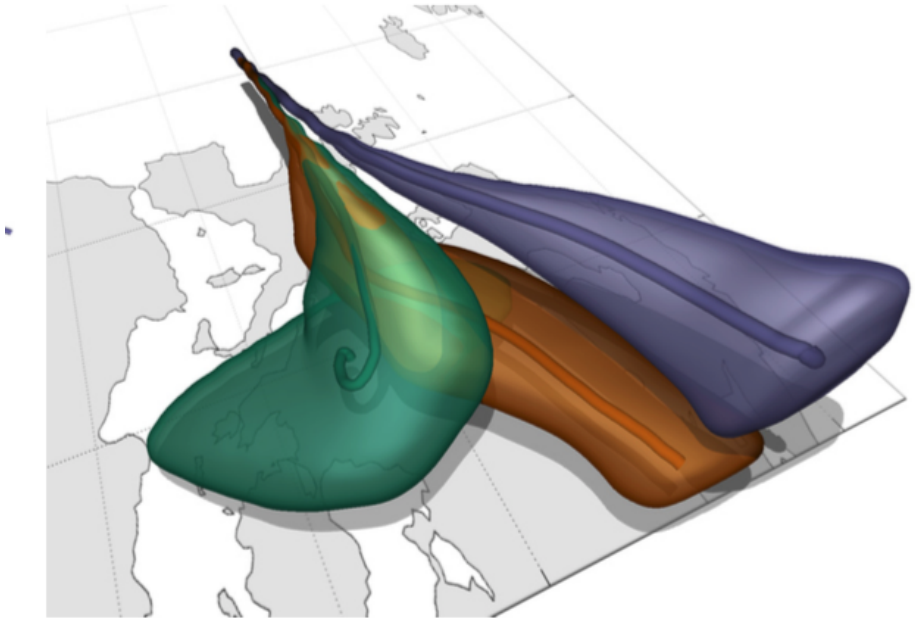
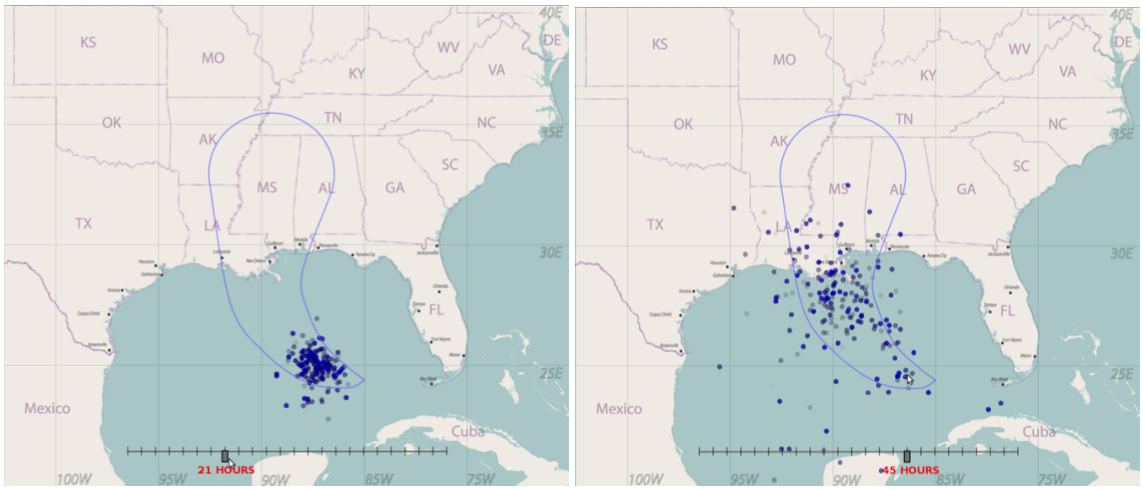


Figure 2.15: Visualization of a hurricane prediction ensemble using the variability plot developed by Ferstl et al. [FBW16]. *Source*: [FBW16].

While these techniques characterize hurricane forecast ensembles extending over time, none of them provides the ability to determine predictions at specific times. To address this limitation, Cox and House [CH13] explored the idea of the interactive visualization of path ensembles at fixed points in time. Rather than rendering complete paths showing a prediction over an entire forecast period, they implemented a time slider to fix a time within the prediction and rendered the corresponding point on each path in the ensemble. Fig. 2.16 shows examples of their approach for Hurricane Katrina 2005 with the time set at 21 and 45 hours from the start of the advisory, comparing them with the corresponding uncertainty cone shown in outline. While this method highlights the uncertainty of the forecast in both space and time, the naive rendering of predicted positions has several limitations. Near the start time of the advisory, data points are tightly clustered, resulting in many overlapping points, making it impossible to take advantage of the point-based display to visually encode other variables using glyphs. At later times in the advisory, there is more spread in the points, but the visual clutter of the display and the overemphasis of the outliers make the position distribution difficult to estimate visually.




(a) 21 hours. (b) 45 hours.

Figure 2.16: Path ensembles sampled at two different times, Hurricane Katrina, 10 AM CDT, August 27, 2005.

## 2.5 Simplicial Depth

The research reported here involves extracting statistical information from ensembles. The recent approaches to do so fall into two main categories, parametric and non-parametric. Parametric methods require an *a priori* assumption of the model describing the data distribution, focusing on estimating the parameters (e.g. mean and variance of a Gaussian distribution) best matching the data. Non-parametric methods attempt to describe the data distribution without an assumption of a model. We use the latter as the hurricane data adopted here has no basis on which to assume a model.

Simplicial depth, which is a widely used and powerful non-parametric approach for describing robust statistical summaries of an ensemble of samples, was developed by Liu [LY90]. As it defines the centrality of an individual point within an ensemble of points, it has the potential to be used to compute a center outward ordering of data. A sample point with a large simplicial depth is considered to be closer to the center of the ensemble and, thus, more representative of the set of points, while a sample point with a small simplicial depth is considered to be less representative. Once the simplicial depth of each point in an ensemble has been determined, the points can be sorted based on their depth, with the indices of the sorted samples providing the structure of a cumulative distribution. Then these indices are normalized by the number of samples to produce a ranking of the points.

Simplicial depth is mathematically represented as follows. Let $V = \{v_0, v_1, v_2, ..., v_{n-1}\}$ be the positions of an ensemble of $n$ 2D points, and let $v_i$, $v_j$, and $v_k$ be three arbitrarily selected members of $V$. Let

$\Delta_{i,j,k}$ denote the triangle formed by these points, meaning the number of triangles is identical to the number of unique combinations of three points from the ensemble which is $\binom{n}{3}$. If $x$ is an arbitrary point in space, we define the indicator function for this point to be

$$I_{\Delta_{i,j,k}}(x) = \begin{cases} 1 & \text{if } x \text{ is inside } \Delta_{i,j,k}, \\ 0 & \text{otherwise.} \end{cases}$$

Then, the simplicial depth of $x$ with respect to data samples $V$ is defined as

$$d(x) = \frac{1}{N_\Delta} \sum_{0 \le i < j < k < n} I_{\Delta_{i,j,k}}(x). \tag{2.1}$$

A straightforward implementation of the simplicial depth calculation for a two-dimensional situation takes $O(n^3)$ computational time. A faster algorithm taking $O(n \log n)$ time has been developed by Rousseeuw and Ruts [RR96].

## 2.6 Radial Basis Function Interpolation

The research reported here involves deriving continuous scalar fields over the spatial region covered by the data samples. To do so, it uses radial basis function interpolation. Radial basis functions (RBF) [BL88] have important applications in several fields requiring scattered data interpolation, most notably in machine learning [O$^+$96] and in computer graphics [CBC$^+$01]. RBF interpolation builds a continuous function from a set of samples using radially symmetric kernel functions of position $x$, of the form $f(x) = \phi(\|x - x_0\|)$, where $x_0$ denotes the kernel center. A number of functions can be used as the radial basis kernel, with one of the most widely used being The Gaussian kernel:

$$\phi(r) = \exp(-\frac{r^2}{2c^2}), \tag{2.2}$$

where $r$ is the distance from the kernel center and $c$ is an adjustable constant controlling the spread of the kernel, similar to standard deviation in a Gaussian distribution. We use the Gaussian kernel because of its property of extrapolating to 0 away from the data and having infinite support, both of which are advantages for handling broadly spread data.

Each data point $i$ is associated with a location $v_i$, a weight $w_i$, and a kernel function $\phi_i$. Then the

RBF interpolation for a given point $x$ is

$$f(x) = \sum_{i=0}^{n-1} w_i \phi_i(\|x - v_i\|). \tag{2.3}$$

If $f_i$ is a scalar value known at each data sample and we impose the condition that $f(x)$ interpolates the available data, each data sample $i$ leads to a linear combination

$$f_i = f(v_i) = \sum_{j=0}^{n-1} w_j \phi_j(\|v_i - v_j\|).$$

Letting $\phi_{i,j} = \phi_j(\|v_i - v_j\|)$ yields the linear system

$$\begin{bmatrix} \phi_{0,0} & \phi_{0,1} & \cdots & \phi_{0,n-1} \\ \phi_{1,0} & \phi_{1,1} & \cdots & \phi_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n-1,0} & \phi_{n-1,1} & \cdots & \phi_{n-1,n-1} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{n-1} \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{n-1} \end{bmatrix},$$

whose unknowns are the weights $w_i$. Letting $\Phi$ represent the matrix, often referred to as the RBF design matrix, $\vec{w}$ the vector of weights, $\vec{f}$ the vector of known sample values, Equation 2.6 can be rewritten as:

$$\Phi \vec{w} = \vec{f}. \tag{2.4}$$

Then the problem is to find a set of weights fitting the available data for the problem space.

A potential issue of applying this technique is that when two samples are nearly identical, $\Phi$ becomes ill-conditioned and results in incorrect solutions. A technique to address this issue is the weighted-decay regularization, which prevents $\Phi$ from being ill-conditioned by adding a constant, $\lambda$, to the diagonal of the design matrix before inverting it. Usually, $\lambda$ is a very small value, e.g. $10^{-4}$, meaning the solution closely approximates the data at the sample points rather than forcing a strict interpolation [ALP14]. All of the RBF work reported here uses a regularization constant of $10^{-4}$.

# Chapter 3

# Visualizing Time-Specific Hurricane Predictions with Uncertainty from a Storm Path Ensemble

This chapter demonstrates an approach for generating and smoothly interpolating robust statistics from path ensembles to produce time-specific visualizations that inherently include uncertainty. As a demonstration, this approach outlines the development of a visualization encoding three levels of positional storm-strike risk for a specific point in time. An example of this visualization is shown in Figure 3.1. In addition to strike position, this method has the potential to be applicable for visualizing such predicted variables as storm speed, wind strength, storm size, and flood risk. The components of this approach that will be of interest to the visualization community include:

- The sampling of each path from the ensemble at a specific time to create an ensemble of points fixed in time
- The application of the concept of simplicial depth to provide a centrality ordering of the time samples
- The development of an adaptive radial basis function interpolation technique that smoothly interpolates simplicial depth
- The design of a geospatial visualization incorporating the concept of risk based on the simplicial depth field

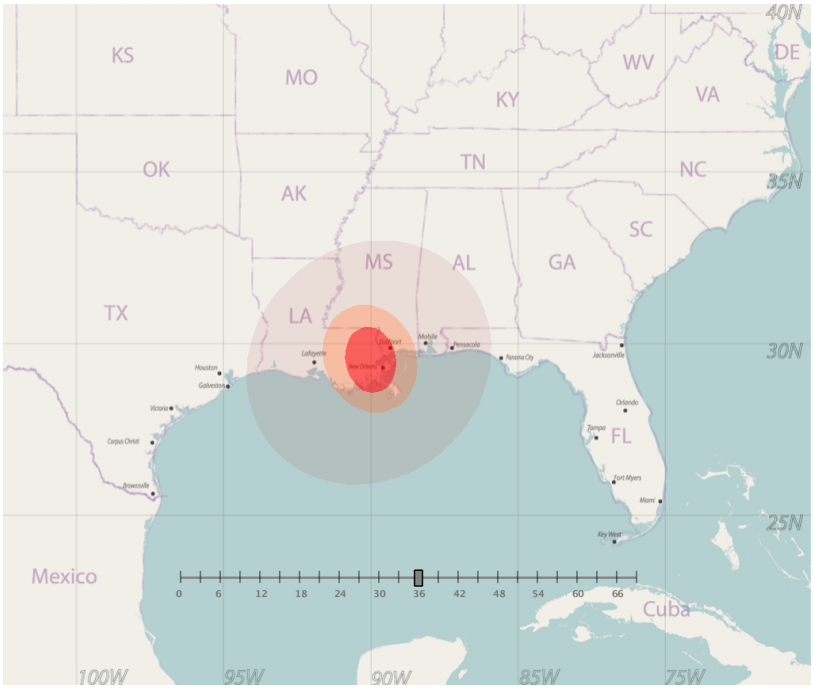



Figure 3.1: Position prediction at 36 hours, NHC Advisory: Hurricane Issac, 1 PM CDT, August 27, 2012.

## 3.1 Methodology

### 3.1.1 Visualizing Simplicial Depth

The approach used here to develop time-specific visualizations from predicted storm path ensembles began by sampling the paths at specific times to obtain an ensemble of locations fixed in time. A method for doing this is to build a simple scatter plot of predicted locations, which tends to create a confusing display as the prediction time increases due to spatial undersampling as seen in Figure 2.16. One possible improvement would be to portray the underlying spatial density distribution implied by the hurricane predictions as a "heat map." However, early attempts of this research to produce heat maps by laying down a spatial grid and counting data points led to displays that were too coarse where data points were tightly clustered and too incoherent were they were widely spread. To address the issues with data density, we use the concept of simplicial depth to give a clean measurement associated directly with a data sample, one not dependent on the local sampling density. Once simplicial depth is assigned to each sample, interpolation methods can be used to create a continuous simplicial depth scalar field from the available samples. This process can be summarized as two steps: computing simplicial depth and interpolating a continuous field.

The first step is to compute the simplicial depth values for all the sample points, which are the pre-

dicted locations from a path ensemble generated to correspond with a storm advisory, using the fast algorithm developed by Rousseeuw and Ruts [RR96]. The simplicial depth of each sample point was subsequently computed, and these sample points were sorted in ascending order based on their simplicial depth. Specifically, if $n$ represents number of samples, the sorted array index of a point divided by $n-1$ is its normalized rank. The set of points within a ranking interval can be visualized by drawing its convex hull. Figure 3.2 shows a scatter plot of rainbow color mapped simplicial depth values, with the red line indicating the convex hull of the $[0\%, 67\%]$ rank interval. Although the rainbow color map is not good for displaying levels because it is

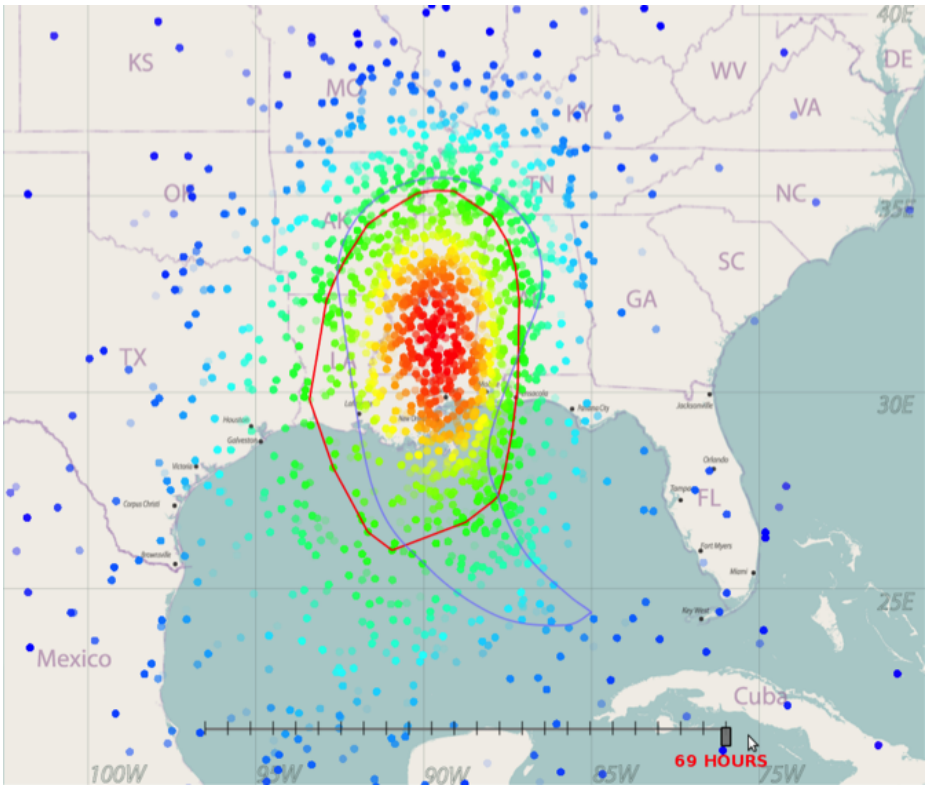



Figure 3.2: A scatter plot of the raw data points. The simplicial depth values are color-coded using the rainbow color map.

not perceptually ordered and thus, is considered misleading for presenting ordered data [BT07], it was used here to follow the NHC's convention for drawing heat maps.

The second step is to interpolate across the computed simplicial depth values in order to provide a smoothly varying continuous representation. One interpolation method is to splat each point onto the map, producing results like those seen in Figure 3.3, where a transparency is applied to each splat proportional to its depth. However, splatting, which depends on sampling to a spatial grid, more importantly leaves many

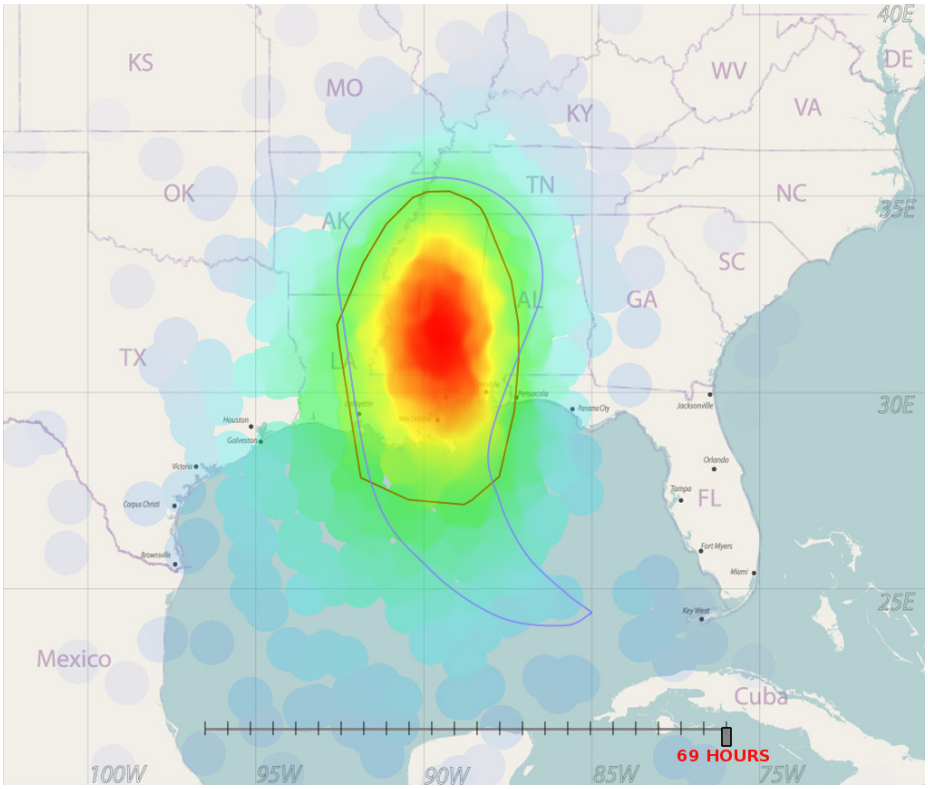



Figure 3.3: Interpolating the raw data points using the splatting algorithm. The simplicial depth values are color-coded using the rainbow color map.

uncolored regions. To address these problems, RBF interpolation can be used as it produces a depth value anywhere in a space, producing smooth visualizations as illustrated in Figure 3.4. Although an improvement

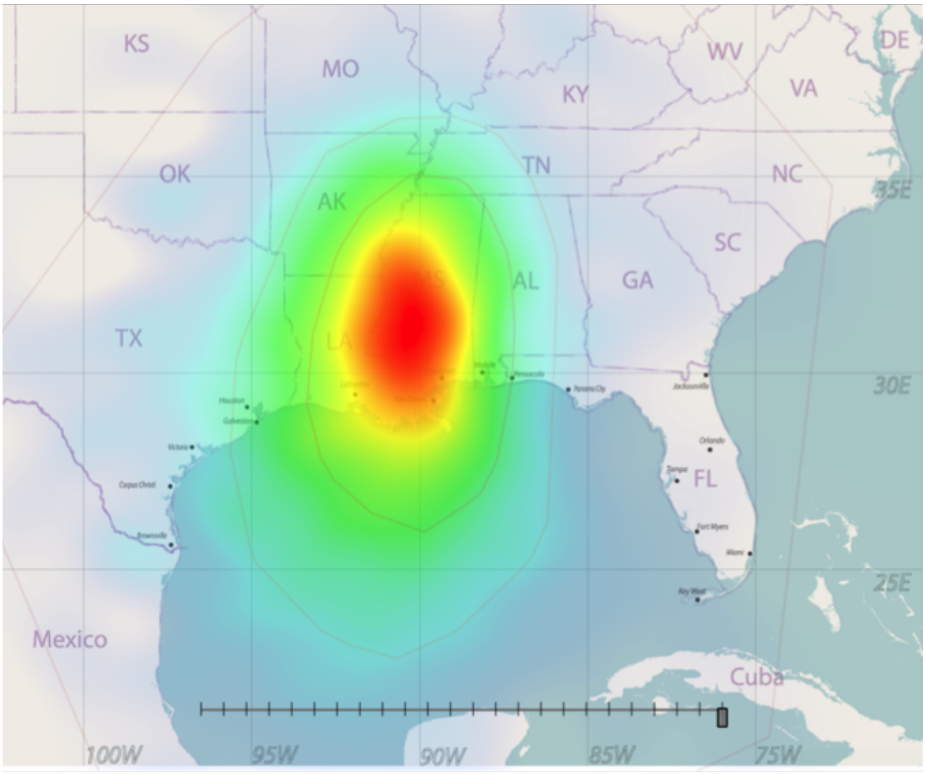



Figure 3.4: Interpolating the raw data points using the RBF interpolation technique. The simplicial depth values are color-coded using the rainbow color map.

over splatting as the central region is smoothly filled, the outer region is highly serrated because of the use of a constant RBF kernel spread parameter with data samples that are unevenly spread. Since the predicted locations are not evenly distributed over the map, an effective interpolation radius for the center region can be too narrow to cover the outer region. Our final interpolation method varies the kernel spread with simplicial density to achieve a smooth interpolation function across the data domain. To address this problem caused by a highly nonuniform data density, this research used an interpolating method that dynamically adjusted the kernel spread as seen in Equation 2.3. The kernel spread parameter is selected based on the prediction density distribution, meaning that dense regions are interpolated with a narrow spread while sparse regions are interpolated with wide spread.

To assign a density value to each sample point, a density field was created using RBF interpolation. To do so, a uniform rectangular grid was constructed over the regional map, and sample points in each cell were counted. Since these grid cells were evenly distributed, the density field was obtained by using RBF

interpolation with a constant kernel spread parameter. This research reported here used the Gaussian kernel centered at $x_0$ as given by $\phi(x) = \exp(\frac{\|x-x_0\|^2}{2c^2})$. The kernel spread parameter $c$ was associated with a bounding box of major dimension $w$ containing all predicted locations, with

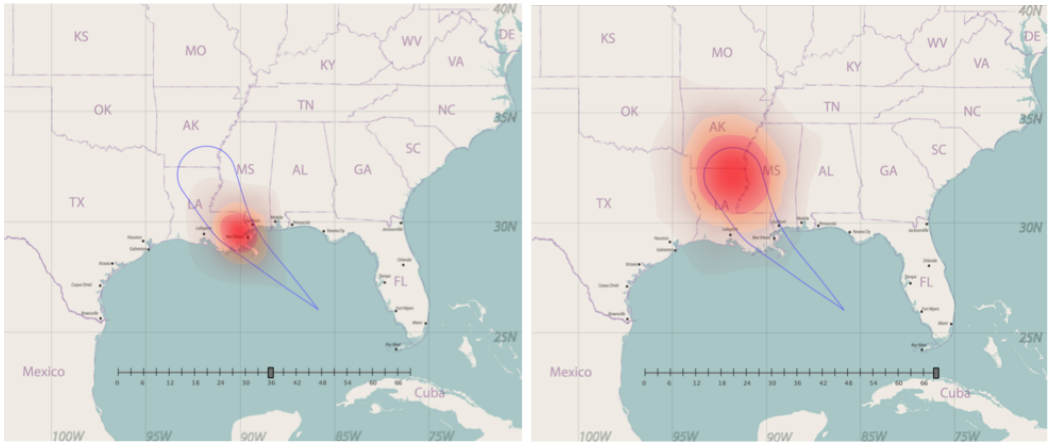
$$c = sw, \tag{3.1}$$

where $0 \leq s \leq 1$ is a user-defined fractional scale factor. Using this kernel spread parameter, all grid cells containing any samples were considered as sample points for building a set of weights to apply to Equation 2.3 for interpolating density.

In addition to simplicial depth $d_i$, for each sample point $v_i$, a density value $\rho_i$ can be computed from the density field and used to determine an appropriate kernel spread parameter $c_i$ for that point. Given the width of an individual grid cell $\delta w$, the kernel spread selected is inversely proportional to the number of data points per unit linear dimension with a constant of proportionality $\lambda$.

$$c_i = \lambda \frac{\delta w}{\sqrt{\rho_i}}. \tag{3.2}$$

These results give a spread parameter that adapts to the density of sample points in the neighborhood of each individual sample point, providing a smooth interpolation across all of the samples. Sample visualizations using this interpolation approach in conjunction with a color encoding meant to clearly show three nested risk regions are shown in Figure 3.5.




(a) 36 hours (b) 69 hours

Figure 3.5: RBF interpolation with dynamically adjustable kernel radius. NHC Advisory: Hurricane Isaac, 1 PM CDT, August 27, 2012.

### 3.1.2 Visualization design

The visualization reported in this work requires the presentation of geolocations including uncertainty. A widely adopted approach for achieving this using a pale (i.e. transparent) blue dot with a radius conforming to some (e.g. 95%) confidence interval has been developed by the GIS and the mobile device communities. Often this blue dot is augmented by a mark indicating its center, and a transparent fade outlines the probability distributions, as exemplified in Figure 3.6. However, a recent series of experiments

Figure 3.6: Two examples of using a transparent blue dot to portray spatial uncertainty.*Source*:[BHMG14]

[BHMG14] provides strong evidence that, contrary to intuition, the uncertainty is more clearly understood without incorporating these augmentations. The visualization design reported here was guided by the premise that including uncertainty in a prediction is of paramount importance. Therefore, the storm position is presented by the three overlapping confidence intervals of 33%, 66% and 99%. These intervals are unembellished except for the use of various colors and transparencies, with the 33% region being the most opaque and the 99% region the most transparent.

The color choices began with the color coding used by the U.S. Homeland Security Advisory System, which employs red, orange and yellow to represent the top three levels of warning, respectively. However, yellow is a poor choice for the application proposed here since highly transparent yellow over a white background is almost invisible. Thus, we replace it with maroon to indicate the cautionary region, while keeping red to indicate the region of highest risk and orange as the medium risk region. Given a depth interval and an associated color, the opacity is given by

$$\alpha = \alpha_0 + d_{\min}\beta, \tag{3.3}$$

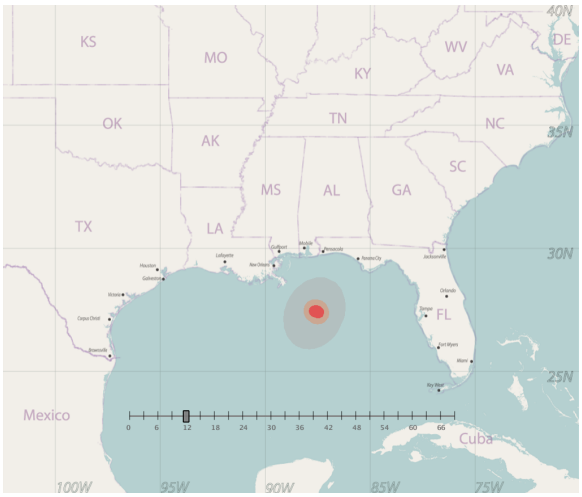
where $d_{\min}$ is the minimum normalized data depth of this interval, $0 \leq \alpha_0 \leq 1$ is the minimum desired opacity, and $0 \leq \beta \leq 1$ is a user-supplied gain. The relevant examples in this work set $\alpha$ to 0.02 and $\beta$ to 0.6.

While the images shown in Figure 3.5 are close to what is envisioned here, the colored depth intervals are more irregular shapes than the standard blue dots. This irregularity, which is induced by the Monte Carlo ensemble generation process, does not provide any useful information. Since the irregular risk regions are nearly elliptical, they were replaced by minimum enclosing ellipses. As the former preserve the aspect ratio of a region along two orthogonal axes, this orthogonality corresponds to the two sources of uncertainty in the prediction: hurricane bearing and speed. Although their effects are not completely independent, speed uncertainty tends to manifest in the elongation of the risk region along the predicted path, while bearing uncertainty tends to broaden the region orthogonal to the path. Figure 3.7 shows three snapshots of a Hurricane Isaac advisory, with the risk regions using the method reported here. The scale factor $s$ in Equation 3.1 is set to 0.35, while the constant of proportionality $\lambda$ in Equation 3.2 is set to 0.3 in this case. The center, the lengths of both the minor and major axes, and the rotation angle of the ellipses, are determined using the image moments-based algorithm proposed by [RVC02].

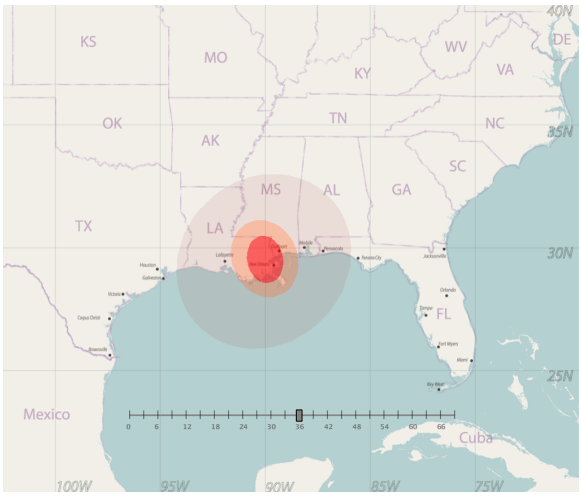
One of the ultimate goals of this research is to develop an interactive application that embeds this approach in a storm position visualization, allowing the user to "scrub" through time. Interpolation over time is easily achieved applying any one of a number of interpolation methods across the known data points forming a path in the ensemble. These points are produced every hour in the NHC ensembles and every three in the method developed by Cox and House [CHL13]. This current work uses simple linear interpolation, after finding that using a higher order method produced visually indiscernible results.

However, an issue producing an animation to simulate scrubbing through time by ellipse representation involves the determination of the major and minor axes, as the ellipses become nearly circular, meaning the computation of these two is not stable, leading to rapid $90^\circ$ flips of ellipse orientation, resulting in disturbing jitters in the animation. Figure 3.8 plots the major axis angle over a series of 138 animation frames for Hurricane Isaac 2012, with the instability being apparent in the top curve of the figure.

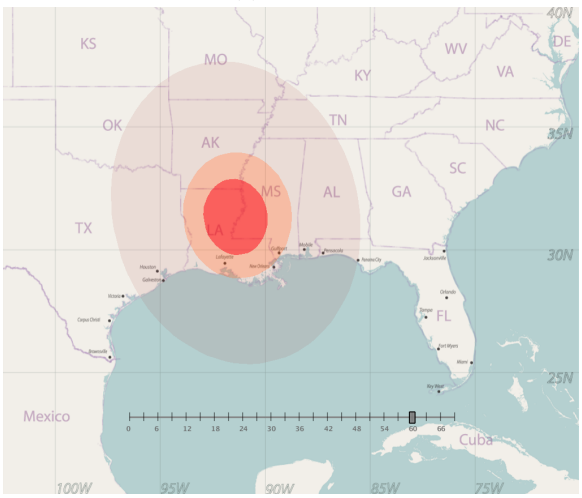
A common way to eliminate the visual noise resulting from this instability is to detect the potential for axis flips using eigenvalue analysis. However, we use a non-linear smoothing filter designed to allow small changes in the angle across time-steps but suppressing large changes. This filter works well and fits more naturally into the signal analysis pipeline. First, the normalized difference between the ellipse orientation angle in time step $i$ and the previous time step $i-1$ is computed using $\Delta\theta = \frac{\theta^{[i]} - \theta^{[i-1]}}{\pi}$.




(a) 12 hours




(b) 36 hours




(c) 60 hours

Figure 3.7: Minimum enclosing ellipses of depth intervals. NHC Advisory: Hurricane Isaac, 1 PM CDT, August 27, 2012.

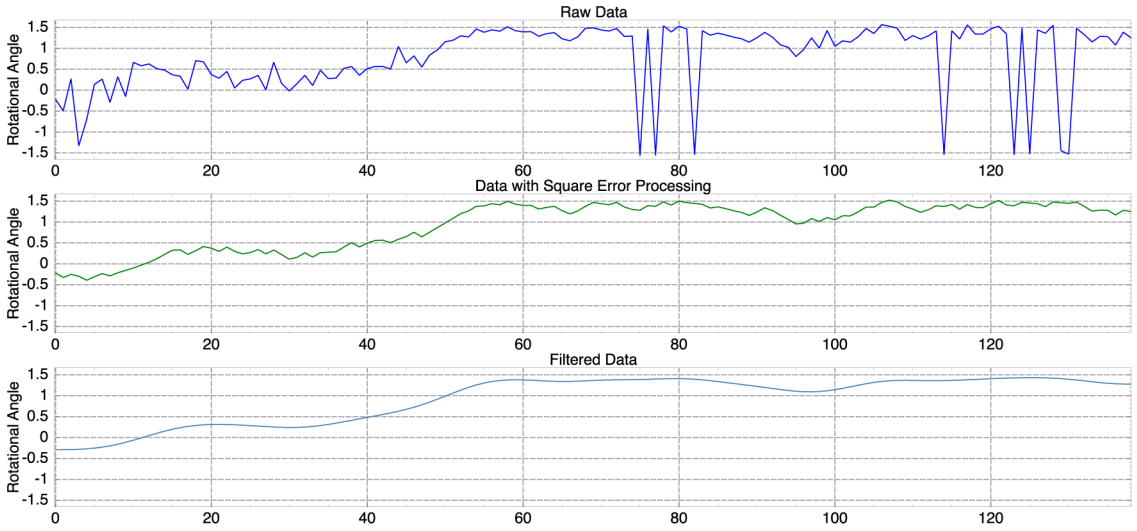



Figure 3.8: Rotation angles as a function of frame number.

Then two weights are computed

$$w_1 = \frac{1}{1+(q\Delta\theta)^2}, \quad w_2 = \frac{(q\Delta\theta)^2}{1+(q\Delta\theta)^2},$$

where $q$ is a parameter controlling the gain of the filter. The filtered angle at the current time is given by

$$\theta^{[i]\prime} = w_1\theta^{[i]} + w_2\theta^{[i-1]}. \tag{3.4}$$

If $\Delta\theta$ is small, $w_1$ dominates, using the current angle, while if $\Delta\theta$ is large, $w_2$ dominates, using the previous angle. In these experiments, setting $q$ to 14 gave the best results. The middle curve in Figure 3.8 shows the result after applying this non-linear filter to the major axis angle shown by the top curve in Figure 3.8.

While the large angular jumps observed in the original curve have been successfully removed, there are still small perturbations that interfere with frame-to-frame visual coherency. To filter these bumps, a filter defined by Equation 3.5, with a kernel of width 5 centered on the current time is applied.

$$\theta^{[i]\prime\prime} = \frac{\sum_{i=-2}^{2}\mu_i\theta^{[i]\prime}}{\sum_{i=-2}^{2}\mu_i}, \tag{3.5}$$

where the $\mu_i$ are the filter weights. These weights were chosen to be nearly Gaussian, by matching the binomial distribution of 6: $\mu_{-2} = \mu_2 = 6$, $\mu_{-1} = \mu_1 = 15$, and $\mu_0 = 20$. The filtered results providing smooth frame-to-frame transitions are shown in the bottom curve in Figure 3.8.

A possible issue with this approach is that the filtered result may not represent the original data accurately. However, the combination of the two filters is applied only to the orientation angle of the ellipse, not to the radii of the major and minor axes. In addition, the nonlinear filter is not strongly sensitive to small angle changes and, thus, removes only large orientation flips in the sequence of visualizations, while the smoothing filter removes only small perturbations from the data, thus removing jitters. As a result, these filters have a negligible effect on the overall orientation of the ellipse angle as can be seen by comparing the curves in Figure 3.8.

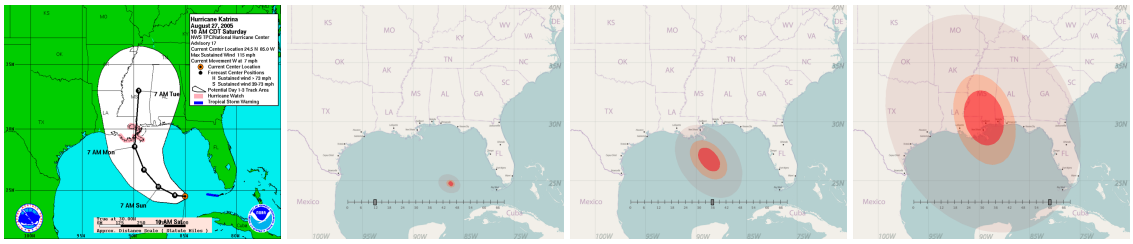
## 3.2 Results

The results obtained from applying the proposed visualization technique on ensembles produced by the National Hurricane Center are shown for a Hurricane Isaac advisory in Figure 3.7. In addition, the proposed method was also applied to advisories for hurricanes Katrina 2005, Alex 2010, Rita 2005, and Ida 2009, using path ensembles generated using the method developed by Cox et al. [CHL13]. The results are shown in Figure 3.9. The system parameters used to generate these results were $s = 0.35$ and $\lambda = 0.15$. The visualizations in the first column show the NHC uncertainty cone for each hurricane, and those in the next three show the predictions at 12, 36 and 60 hours.

Even though the approach reported here involves six tunable parameters, five did not need to be adjusted across different storm advisories. The only parameter needing adjustment across ensemble generation methods was $\lambda$, which controls the adaptive kernel spread parameter in Equation 3.2. Because the method developed Cox et al. [CHL13] generates a broader spread of hurricane paths than the method used by the NHC, the density in the denominator of Equation 3.2 tends to be low, meaning the value of $\lambda$ is decreased to compensate. Thus, the method reported here appears to be robust, requiring only the tuning of one parameter depending on the ensemble generation method.

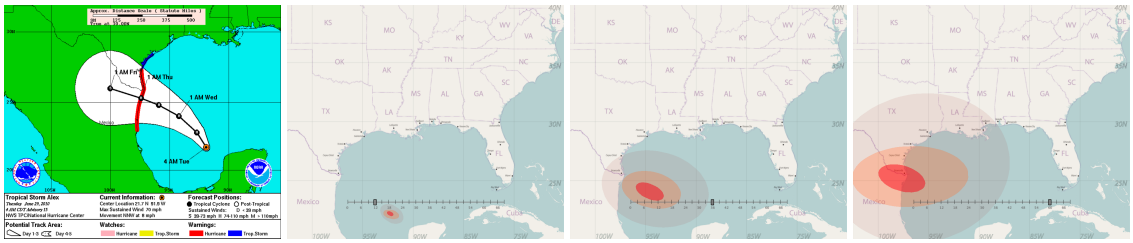
Even though a criticism of this visualization is that as time progresses in a prediction, the size of the risk regions increases, leading to the perception that the storm itself is increasing in size. This problem is also inherent in the NHC uncertainty cone as well as in geospatial displays that attempt to track dispersion in a prediction using a summary display. This problem is difficult to solve, as long as spatial extent is being used as an indication of the spatial uncertainty. As the current research provides strong evidence that ensemble displays do not induce this same perceptual anomaly [RBP$^+$16, CG15, CHL13], the next goal of this research is to extend the work reported here by representatively resampling the forecast ensemble to produce a set of

Hurricane Katrina, NHC advisory 10 AM CDT, August 27, 2005



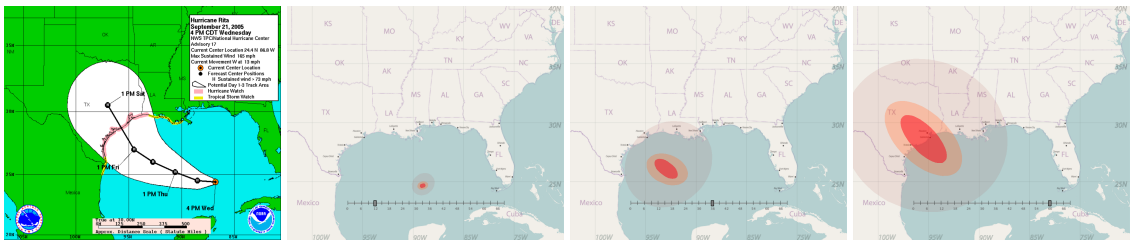
uncertainty cone 12 hours 36 hours 60 hours

Hurricane Alex, NHC advisory 4 AM CDT, June 29, 2010



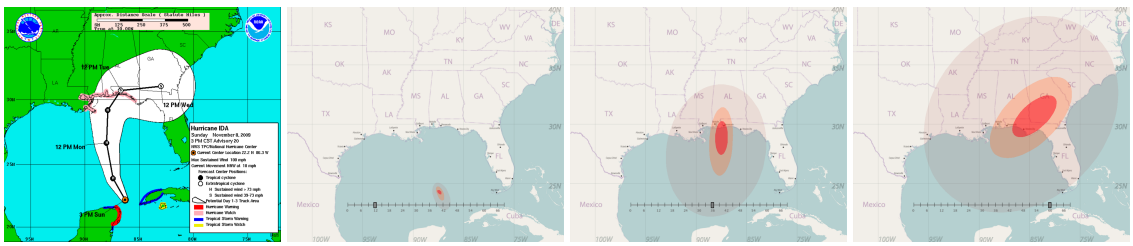
uncertainty cone 12 hours 36 hours 60 hours

Hurricane Rita, NHC advisory 4 PM CDT, September 21, 2005



uncertainty cone 12 hours 36 hours 60 hours

Hurricane Ida, NHC advisory 3 PM CST, November 8, 2009



uncertainty cone 12 hours 36 hours 60 hours

Figure 3.9: Time-specific visualizations of risk regions for four hurricane advisories.

exemplar storm positions that can be displayed as points, but without the visual clutter and confusion of the early work conducted by Cox and House shown in Figure 2.16.

# Chapter 4

# Uncertainty Visualization by Representative Sampling from Prediction Ensembles

Direct ensemble displays have experimentally been shown to be superior to summary displays in terms of both more accurately conveying the spatial distributions associated with prediction uncertainty and minimizing the confusion concerning the spatial attributes of a prediction with its uncertainty [CHL13, RBP+16]. An additional potential advantage of direct ensemble displays is that they reduce the dimensionality of the display elements. Instead of an ensemble of points being represented with a summary area or volumetric display, the points remain points in 2D or 3D. Likewise, paths remain line segments. These advantages of point displays mean that a glyph can be displayed at each point to convey additional information about the ensemble member. For example, in a hurricane position display, each ensemble member could be represented by a glyph encoding the hurricane category.

While directly displaying the ensemble used to make a prediction has several advantages, it also has visualization problems, the most important being the potential for it to become a confused jumble. One widely adopted sampling approach to avoid this problem is through blue noise sampling, which generates a set of samples randomly located but remaining spatially well separated. For instance, Wei [Wei10] developed a multi-class blue noise sampling technique, and more recently, Chen et al. [CCM+14] utilized this technique to develop a visualization framework for multi-class scatter-plot data. This sampling approach detects collisions

of points using a matrix encoding the inter-class and the intra-class minimum distances in such a way that both the mixed samples and the samples of individual classes are uniformly distributed. This approach can be extended to support adaptive sampling by designing spatially varying functions and constructing a distance matrix at each sample using these functions. Although using this approach the resulting subsampled distributions look similar to the originals, there is no guarantee, however, that they are nearly identical. For the research reported here, control of these distributions is essential as the goal is to use them to communicate implicitly the uncertainty in the prediction.

This work aims to develop approaches to make effective displays from prediction ensembles that are a set of points. It develops a novel methodological framework for either selecting a subset from the original ensemble or resampling to create a new set that supports a variety of visualization approaches. This resulting subset must have the following properties:

- It must correctly convey the key statistical spatial properties of the original ensemble.
- It must be representative of the full ensemble if used as a basis for scattered data interpolation.
- The points must be spatially well separated to minimize occlusion.

First, a space in which ensemble members are uniformly distributed is constructed, and then selection or resampling processes are conducted in this space before projecting the selected points back to the original space for display. Figure 4.1 shows an example of this methodology applied to an ensemble of predicted hurricane paths from the NHC. The original ensemble (Figure 4.1a) is sampled 36 hours into the prediction (Figure 4.1b) to show possible hurricane positions in the prediction data. These are displayed using glyphs by the NHC to indicate hurricane geospatial position and strength on its maps. A spatial warp is constructed (Figure 4.1c) to transform this set of geospatial samples into a space where they are uniformly distributed (Figure 4.1d) referred to as the UD space. Subsequently, a subset of the samples is selected (Figure 4.1e) to achieve a Poisson Disk distribution in UD space. Finally, the selected points are transformed back into the original space, abbreviated OD space, and displayed geospatially using the NHC glyphs (Figure 4.1f).

Even though a lot of algorithms have been developed for selecting exemplars, directly applying them to compute a small collection of representatives from the forecast ensemble does not meet our requirements. For instance, the Orthogonal-Least-Squares (OLS) [CCG91] forward selection algorithm extracts a subset of samples by minimizing the global interpolation error. The obtained subset may be effective for interpolation purposes, but it does not guarantee preservation of the original distribution and has no mechanism to maintain spatial separations between samples. In addition, minimizing the interpolation error implies ignoring the

outliers included in the original ensemble, which are important to be conveyed in the visualization of the prediction of any severe weather event. Another important work for selecting representatives was conducted by Ehsan et al. [ESV12], which finds a subset from a dataset by minimizing the pair-wise dissimilarities. This subset is effective for distinguishing clusters included in the original dataset, but has similar limitations to those that the OLS approach has. While these algorithms cannot be directly utilized in the original display space in our application, they can be potential alternatives to selecting the desired subset in the UD space.

The primary contributions of the work reported here include:

- Development of a sampling technique that extracts a representative subset of an ensemble of points that not only accurately preserves the original distribution but also maintains good spatial separation among the subset members.
- Development of two dynamic visualizations of a time-sampled hurricane prediction ensemble. One of these conveys the hurricane intensity using glyphs at the sample points, while the other conveys storm size using circles the radii of which are determined by the storm size.
- Improvement of a previous summary display developed in Chapter 3, showing how this sampling technique can be used to support smooth and accurate interpolation.
- A cognitive experiment comparing participants' interpretations of ensemble visualizations to those of existing summary displays.

## 4.1 Selection Algorithms

This section provides a brief introduction to the mathematical and algorithmic foundation for the approach used here for finding representative subsets from point ensembles. First, this study involves an intermediate process that computes a continuous density field from a set of discrete data samples. The RBF interpolation technique was utilized here to achieve this goal. Details about this technique have been fully discussed in Section 2.6. In addition, we use two mechanisms to select a representative subset from originals in the UD space: the orthogonal least squares algorithm and the weighted sample elimination algorithm.

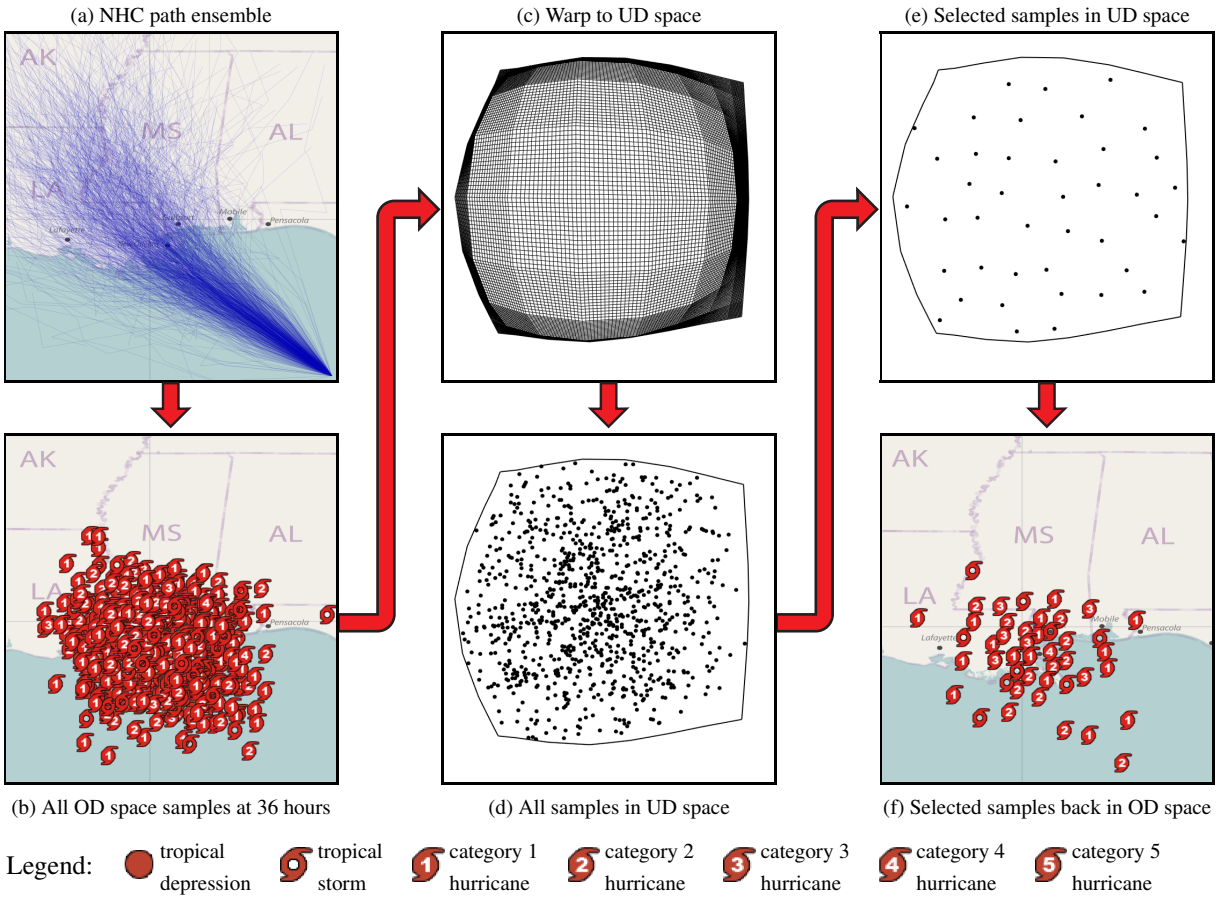



Figure 4.1: The pipeline for sampling from a path ensemble to produce a well structured time-specific ensemble visualization.

### 4.1.1 Subset Selection via Orthogonal Least Squares

Simply using RBF interpolation to form a field from a large set of point samples is problematic for several reasons. Most importantly, each of the point samples involves uncertainty, meaning fitting a surface exactly to these point samples will result in an overfitting that will not generalize the data based on its statistics. Second, for a set of $M$ samples, the RBF design matrix $\Phi$ will be of size $M \times M$ and will be dense unless kernel functions of small finite support are used. Thus, finding a solution to the linear system for the weight vector $\vec{w}$ will be slow. Finally, for samples from a large ensemble will tend to be located very near to one another, resulting in ill-conditioning of the matrix. The forward selection algorithm [O$^+$96] is an approach for addressing RBF interpolation problems by selecting a small representative subset to use as a reduced basis for interpolation [CCA96, CCG91]. OLS begins with an empty subset to which one-point samples are added one at a time; each selection of these samples is made to minimize the sum squared RBF interpolation error (SSE) in relation to the original sample points. This process can be accelerated by using Orthogonal Least Squares (OLS) method [CCG91], meaning each selection of samples is orthogonal to the samples already selected.

### 4.1.2 Subset Selection via Weighted Sample Elimination

A completely different approach for choosing a subset of point samples from an original set is to optimize for spatial layout rather than the mean squared interpolation error. Choosing a subset of an ensemble that maintains the Poisson Disk property (i.e. that all randomly located points are tightly packed but lie beyond a chosen Poisson Disk radius $r$ from one another) ensures that the space under consideration is being sampled in an optimal way, similar to what is typically used in signal processing theory [Coo86]. A number of techniques have been developed to generate Poisson Disk samplings over a spatial domain, including acceleration methods using spatial data structures [DH06] and varying the disk radius over a domain using importance measures [MF92].

Recently, Yuksel [Yuk15] developed the weighted sample elimination approach, which begins with an existing sample set and returns a radius and a subset of a specified size that exhibits the Poisson Disk property for this radius. This approach is suitable for the purpose of this study, which is to determine a representative subset of a spatially distributed ensemble. This approach uses a kd-tree structure to determine a set of neighbors for each point and a heap data structure organized by a sum of weights representing the accumulated distances of each point from its neighbors. The algorithm iteratively eliminates the point with

the highest weight, recomputes the summed weights, and reorganizes the heap until the desired number of samples is reached. The Poisson Disk radius is then the minimum distance between the remaining points.

## 4.2 Methodology

The steps of our approach for selecting a representative subset of $N$ points from an ensemble of $M$ points are given below:

1. *Determine the bounding region of the M points in OD space.*

   The remaining steps are completed with respect to the bounding region of the ensemble points in the original density, or OD, space.

2. *Compute a continuous density field over the bounding region.*

   Within this bounding region, ensemble points are used to estimate a continuous density field, as explained more fully below in Section 4.2.1.

3. *Construct a warp from OD to UD space.*

   This field is used to construct a warping function that maps the points in OD space to a space in which the density per unit area is constant, i.e. the uniform density or UD space. The construction of this mapping function is discussed in Section 4.2.2.

4. *Map each point from OD space to UD space.*

   This warping function is used to map each point $(x_i, y_i)$ in OD space to $(u_i, v_i)$ in UD space, such that samples in UD space are uniformly distributed as explained more fully in Section 4.2.2.

5. *Select a set of $N < M$ points in UD space.*

   A point selection algorithm is then used in UD space to select a subset of samples as explained in Section 4.2.3.

6. *Project the N selected points back to OD space for display.*

   Finally, the selected points in UD space are mapped back to OD space for display as discussed in Section 4.2.4.

The selected subset being uniformly distributed in UD space guarantees that the corresponding sample set in OD space is representative of the statistical properties of the original ensemble.

### 4.2.1 Computing the Continuous Density Field

The density field is computed over the sample space using RBF interpolation after assigning a local data density value to each sample point. A grid-based discretization for the density estimation was not used here because the input data from the original ensemble may be distributed in a non-uniform way, resulting in sampling issues. Instead, the samples of the original ensemble utilizing a $k$ nearest neighbors ($k$NN) approach [Alt92] was used.

Given a sample point $i$ with position $\vec{x}_i = (x_i, y_i)$ from a data set of size $M$, the $k$NN density $\sigma_i$ at this point is defined as

$$\sigma_i = \frac{k}{M\pi{\rho_i}^2}, \tag{4.1}$$

where $\rho_i$ is the radius of the circle with its center at $\vec{x}_i$ that minimally encloses the $k$ nearest neighbors of the point $i$. A kd-tree was used for rapid determination of the $k$ nearest neighbors of each sample point.

However, interpolating the density field from all of the points in the ensemble would result in overfitting. Therefore, the forward selection implementation of the OLS algorithm [O$^+$96] was used here to select a subset that minimizes the mean squared interpolation error. For each sample point $i$ an RBF spread parameter that adapts to the local density measure was computed using the following equation

$$r_i = \beta\frac{w}{\sqrt{\sigma_i}}, \tag{4.2}$$

where $w$ is the largest dimension of the bounding region in OD space and $\beta$ is a user settable constant. As a result, the RBF interpolation algorithm can interpolate scattered data with widely varying densities.

It is possible to use the subset of the original points obtained with this procedure as the subset for displaying the ensemble as well. However, by design this subset is more uniformly distributed than the original data set. Therefore, it does not provide a good representation of the statistical properties of the original ensemble. Furthermore, this subset can also contain points that are too close together, resulting in glyph occlusion. These two problems can be seen in Figure 4.2. Comparing the original ensemble of time-specific hurricane position predictions (Figure 4.2a) to the subset of 100 sample points chosen by the OLS algorithm (Figure 4.2b), it is apparent that the resulting subset is too uniform and that some points are too close together. To avoid this problem, the UD space was used for the sample selection.

(a) entire set (b) 100 selected samples

Figure 4.2: An optimal subset chosen using OLS in the original space that is too uniformly distributed.

### 4.2.2 Mapping from OD to UD Space

There are a number of ways to construct the mapping seen below:

$$OD, UD \subseteq \mathbb{R}^2, \;\; f : UD \to OD, \;\; (x,y) \mapsto (u,v). \tag{4.3}$$

However, not all of these result in a warp suited for the purpose of this study. The warp function should minimize both shear and non-uniform scaling since the selection process in UD space will be based on Euclidean distances between the points. Locally, the mapping should be as close as possible to uniform scaling.

It is possible to define a mapping from a non-uniform distribution to a uniform distribution using a cumulative density function. In $\mathbb{R}^2$, the probability density $p$ at position $(x,y)$ can be written as

$$p(x,y) = p_x(x)p_y(y|x), \text{ where} \tag{4.4}$$

$$p_x(x) = \int_{-\infty}^{y} p(x,y)dy \;\text{ and }\; p_y(y|x) = p(x,y)/p_x(x), \text{ yielding} \tag{4.5}$$

$$u(x,y) = \int_{-\infty}^{x} p_x(x)dx, \text{ and } v(x,y) = \int_{-\infty}^{y} p_y(y|x)dy. \tag{4.6}$$

Although this mapping results in a uniform distribution, it introduces significant shear along the $v$ axis such that points that are distant from one another in OD space can be moved near to one another in UD space and vice versa. Therefore, the Euclidian distance between points in UD space is not a good metric for selecting a subset.

The approach for addressing this issue that is used here is a Gauss-Seidel style relaxation process which begins by splitting the OD space into a number of grid cells of uniform sizes. In UD space, this grid is deformed such that the area of each grid cell corresponds approximately to the density within the grid cell in OD space. The deformation of this grid represents the warping function. A point $(x_i, y_i)$ in a grid cell in OD space is mapped to $(u_i, v_i)$ in the corresponding cell in UD space, using matching barycentric coordinates.

#### 4.2.2.1 Relaxation for the Grid-based Warping Function

Without loss of generality, let us assume that the initial grid is a square with dimensions $S \times S$. Since UD space has a uniform density, each deformed grid cell in UD space should have the same average density

$$d_a = \frac{\sum_{j=1}^{S^2} d_j}{S^2 A}, \tag{4.7}$$

where $d_j$ is the initial density of the cell $j$ and $A$ is the initial area of a grid cell. In practice, the initial density $d_j$ of a grid cell is estimated by sampling at the center of the cell using the $k$NN density field described in Section 4.2.1. The target area of a cell $A_j$ can be written as

$$A_j = A\frac{d_j}{d_a}. \tag{4.8}$$

For computing the deformed grid, a target length is assigned to each edge of the grid using the $A_j$ values of the neighboring cells. The target length of an edge $k$ on the perimeter of the grid that has a single cell neighbor $j$ is

$$L_k = L\sqrt{\frac{A_j}{A}}, \tag{4.9}$$

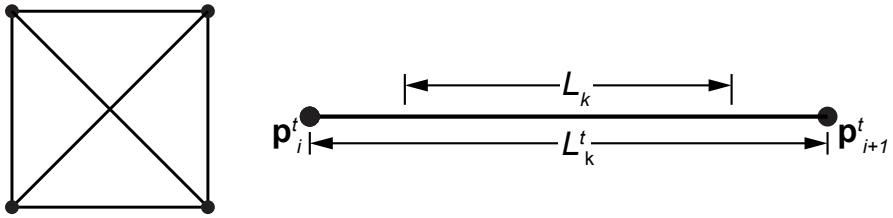
where $L$ is the initial length of a grid edge. For other edges shared between two cells, the average of the two target length values computed for each face is used. Two diagonal edges for each cell are constructed to

minimize shear during the relaxation. The target lengths of the diagonal edges are computed similarly, using the target areas of their cells. The structure of a single cell is shown in Figure 4.3a.

This relaxation process deforms the length of each edge $k$ towards the target length $L_k$ at each step $t$ by moving its vertices along the edge direction. Let $\vec{p}_i^{\,t}$ and $\vec{p}_{i+1}^{\,t}$ be the positions of the two vertices of an edge $k$ at relaxation step $t$ as shown in Figure 4.3b. The updated position of the vertex $\vec{p}_i^{\,t+1}$ due to the update operation for edge $k$ can be written as

$$\vec{p}_i^{\,t+1} = \vec{p}_i^{\,t} + \alpha \left( \frac{L_k^t - L_k}{2} \right) \left( \frac{\vec{p}_{i+1}^{\,t} - \vec{p}_i^{\,t}}{L_k^t} \right) , \tag{4.10}$$

where $L_k^t = \left| \vec{p}_{i+1}^{\,t} - \vec{p}_i^{\,t} \right|$ is the length of the edge at relaxation step $t$ and $0 < \alpha \leq 1$ is a user adjustable acceleration factor such that if $\alpha = 1$, then $L_k^{t+1} = L_k$ at the end of the update for edge $k$.

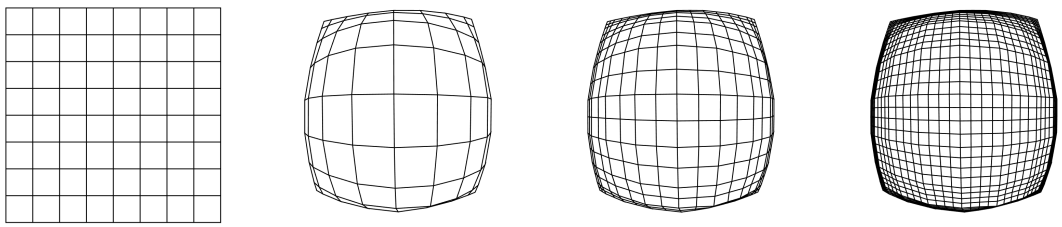



(a) cell model (b) edge model

Figure 4.3: Model used for a grid cell and its edges

#### 4.2.2.2 Hierarchical Relaxation

To achieve a high quality warping function requires a high-resolution grid. However, when the displacement of the grid vertices due to its deformation is much larger than the length of a grid edge, the



(a) OD domain (b) UD codomain (c) UD codomain (d) UD codomain

level 0 level 1 level 2 level 3

Figure 4.4: Hierarchical refinement for computing the warping function

relaxation process can result in crossed edges, inverting cells and thus changing the grid's topology, leading to instability. To avoid these issues, the process used here involves a hierarchical progressive-refinement approach as illustrated in Figure 4.4. Relaxation is used to compute a low-resolution deformed grid. Then, the cells are subdivided, and the relaxation process continues for the subdivided grid with higher resolution. Using this approach, large deformations of the grid are resolved using lower resolution grids, and fine details of the deformation are introduced using the relaxation of higher resolution grids.

The relaxation process used here to obtain a solution at any level in the hierarchy has two stopping conditions. First, the sum of squared errors (SSE), defined as the sum of squared differences between desired areas and current areas of all grid cells, is monitored:

$$\mathrm{SSE} = \sum_{p=1}^{S_k^2} (A_p^D - A_p)^2 \tag{4.11}$$

where $S_k$ denotes the grid dimension at the $k$th level in the hierarchy. The relaxation is terminated when the SSE is less than a small threshold. The second stopping condition is the detection of an inverted cell. When a target cell area is very small, the relaxation process can still produce inversion. If an inverted cell is detected, the system is restored to its previous state, and the relaxation terminates at that level of the hierarchy and continues at the next lower level.

At each stage of the refinement process, the area of each grid cell is adjusted so that its density equals the average density of the OD domain. In the implementation used here, the grid dimensions are set as powers of two, defining the finest grid dimensions $S \times S$ with $M$ samples using

$$S = \lfloor \log_2 \gamma M \rfloor \tag{4.12}$$

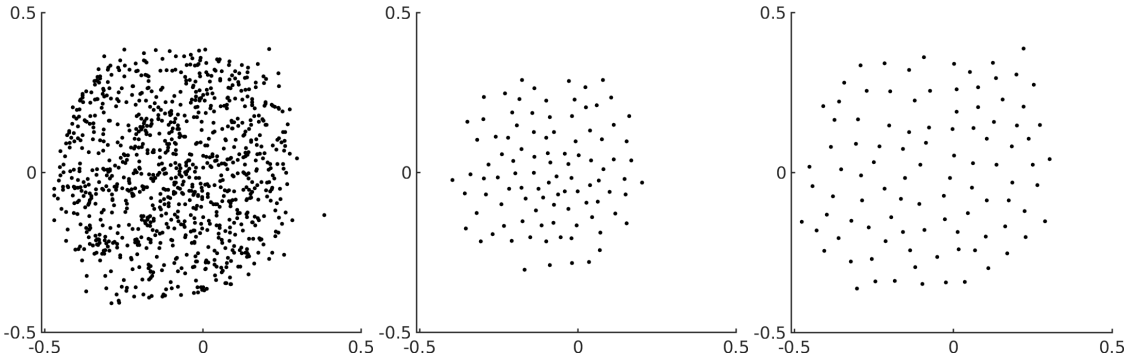
where $\gamma$ is a user adjustable fraction between 0 and 1. If $S_0$ is the dimension of the grid at the coarsest level, this method requires

$$T = 1 + \log_2 S - \log_2 S_0 \tag{4.13}$$

hierarchical levels to compute the deformed grid of the highest resolution.

### 4.2.3 Selecting a Representative Subset

To select a representative subset, first all samples are transformed from the original OD space to UD space. Subsequently, a representative subset of the samples is selected in UD space using either the




(a) Full ensemble (b) Subset selected via OLS (c) Subset selected via WSE

Figure 4.5: Comparison of samples selected via OLS and WSE algorithm in UD space

Orthogonal Least Squares (OLS) algorithm explained in Section 4.1.1 or the Weighted Sample Elimination (WSE) method presented in Section 4.1.2.

These two methods select subsets with different features. The OLS method is a selection approach based on a construction of an RBF system. To avoid the high computational cost of the naive OLS algorithm, we use Orr's forward selection algorithm [O$^+$96], which monitors the reduction in the SSE of the RBF system as it selects the minimal subset satisfying a specific error criterion. While building the RBF, the right side of the linear system is assigned by computing the simplicial depth values of the original samples. A sample with a larger simplicial depth value is considered to be more representative and, thus, will tend to have a higher chance to be selected by the OLS algorithm. Figure 4.5a shows the entire ensemble, and Figure 4.5b shows the subset selected by the OLS algorithm. This subset is concentrated in the central region of UD space and, thus, is tightly clustered in OD space.

On the other hand, the WSE algorithm attempts to select a subset that maintains the Poisson Disk property. Therefore, the samples in the subset selected are uniformly spread over the entire UD space and, thus, are more broadly spread in OD space as can be seen in Figure 4.5c. In the WSE algorithm, the sum of the weights of neighbors of a sample determines the sample depth in a heap data structure. Because samples near the boundary of UD space have fewer neighbors, their weights are biased, making them more likely to be kept in the samples selected. To address this problem, the concave hull of the collection of points is extracted, and extra samples are randomly generated in the region outside of this hull so that the density of the boundary region matches the average density of UD space. These generated samples are considered when computing the weights of the original samples but are never selected by the algorithm.

### 4.2.4 Inverse Mapping from UD to OD Space

The inverse mapping of the original points from UD space to OD space maps them back to their original coordinates in OD space. However, if it is desired to resample in UD space to create new points, an inverse mapping is required to find the corresponding positions in OD space. This inverse mapping is defined similarly to the forward mapping but begins with the deformed grid. Given a point $(u_i, v_i)$ in UD space, first the deformed grid cell that contains this point is found. Since the relaxation procedure used here ensures that no grid cell is either inverted or has crossing edges, there exists a unique deformed grid cell for any point in UD space. A space partitioning structure is utilized here for quickly finding the corresponding cell. Then, barycentric coordinates are used to project the point to the corresponding undeformed OD space cell.

## 4.3 Validation and Visualization Design

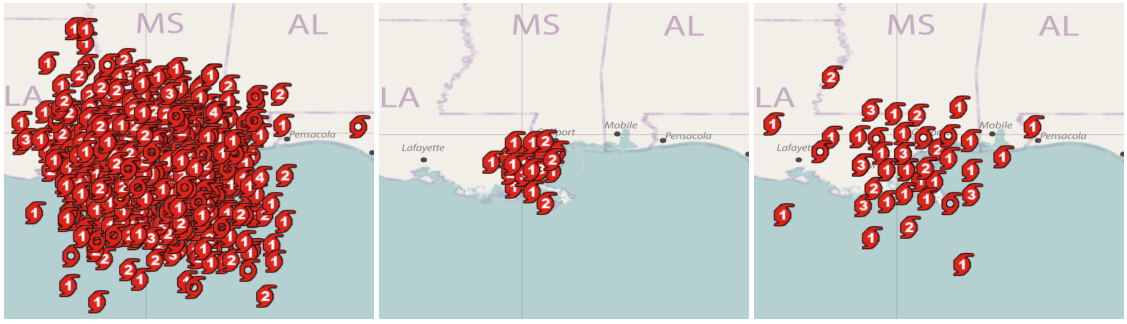
This section describes the experimental results demonstrating the utility of the technique reported here for selecting a representative subset of time-specific hurricane positions from a prediction path ensemble. It first demonstrates that these subsets closely reproduce the spatial distribution encoded by the full point ensemble. Then it shows that this approach can be used to support two forms of representative visualization superimposed on a map: a point ensemble display using glyphs to show a set of predicted storm locations annotated with additional storm information, and a summary display showing three levels of risk regions, similar to the displays developed in Chapter 4.

This approach is exemplified using the NHC's ensemble of 1,000 predicted paths for Hurricane Isaac, a prediction made approximately 36 hours before the hurricane made landfall in Louisiana on August 29, 2012. Since the paths are sampled in time and encoded with predicted storm characteristics such as wind speed and storm size at regular time intervals, they can be used to produce spatio-temporal visualizations. We have already shown an example of our complete process using this NHC prediction in Figure 4.1.

There are five user-defined parameters in this approach. To determine the number of nearest neighbors for the $k$NN density calculation in Equation 4.1, we let $k = \lfloor 0.01M \rfloor$, where $M$ is the data set size. To determine the radial basis function spread for the $k$NN density calculation in Equation 4.2, we set the scaling parameter $\beta = 0.01$. To determine $S$, the grid dimension at the finest level of the hierarchy in Equation 4.12, we set $\gamma = 0.2$. Thus, given 1,000 predicted hurricane locations, the final grid resolution is $S = 128$. The coarsest grid resolution was selected to be $S_0 = 8$, meaning there are $T = 5$ levels in the hierarchy (Equation 4.13). The acceleration constant used in performing the relaxation algorithm defined by Equation 4.10 was

set to $\alpha = 0.066$. For all the resulting demonstration images, the OD space is measured geospatially by latitude and longitude, where the UD space is centered at the origin by dimensions of one unit by one unit: $[-0.5, 0.5]$.

Figure 4.6 shows an example of how the choice of the OLS versus the WSE selection algorithm affects the layout of the samples selected. The sizes of the subsets selected were based on a heuristic that




(a) Full ensemble (b) Subset selected via OLS (c) Subset selected via WSE

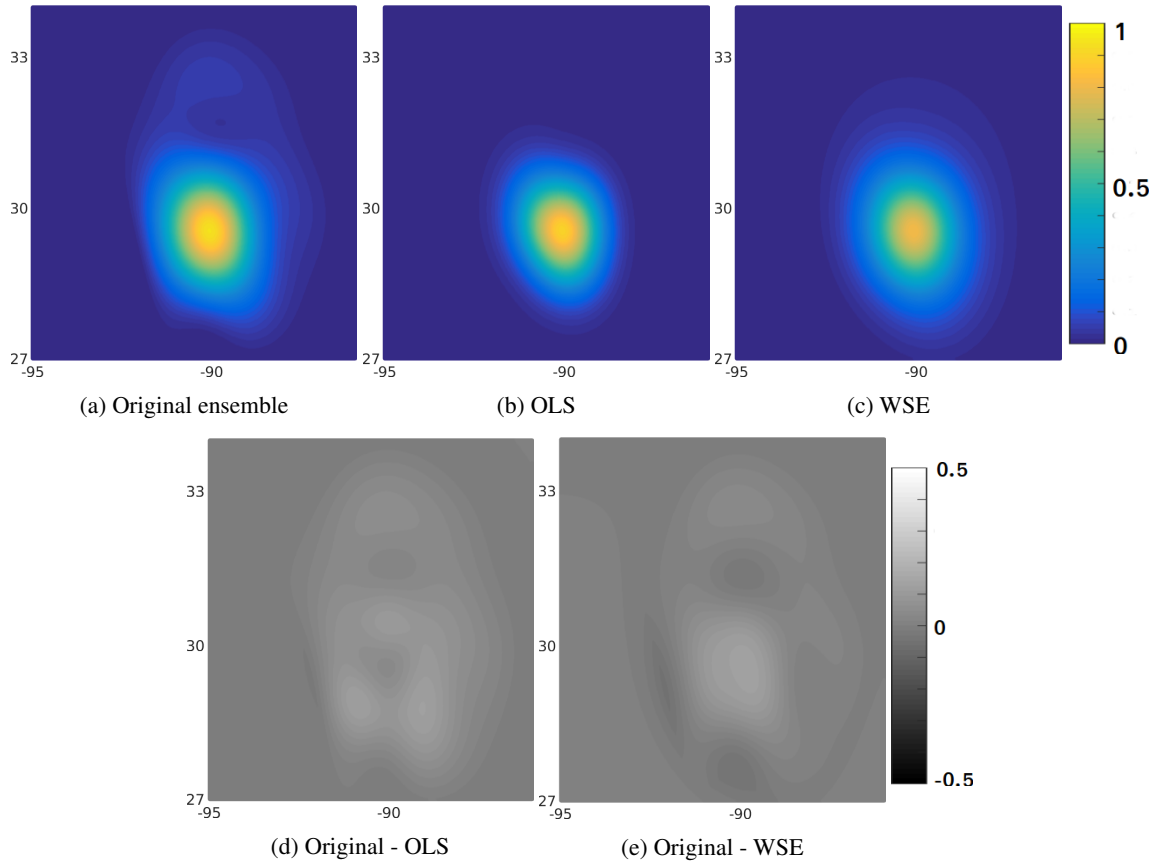
Figure 4.6: Comparison of samples selected via OLS and WSE algorithms, Hurricane Isaac, 2012, prediction at 36 hours.

specified the tightness of spacing. Given the area of a bounding region $B$ of the original set of samples, the size of the subset is determined as $N = \delta B$, where $\delta$ denotes the desired number of samples per unit area. In the implementation reported here, $\delta$ is set to 0.72 for the OLS algorithm and 1.45 for the WSE algorithm, reflecting the relative tightness of spacing between the two approaches. Comparing Figure 4.6b, which has 21 points, with Figure 4.6c, which has 42, indicates that the OLS algorithm tends to produce a subset that is more concentrated around the most likely position of the hurricane, while the WSE algorithm tends to produce a subset that is more widely spread, indicating not only the most likely storm position but the possibility of outliers. The reasons for these differences have already been discussed in Section 4.2.3.

It is important to note that while the original ensemble in Figure 4.6 appears to spread over an even larger area than both of the subsets generated by the OLS and the WSE, these outlier samples away from the center appear with very low probability, and the density of the ensemble samples near the center is completely imperceivable due to occlusion. Therefore, the visualization of the full ensemble can be misleading.

Even though the two selection algorithms produce subsets with different layouts, both are representative of the original ensemble. One way to verify this is to test if a selected subset can be used to reproduce a distribution supported by the original ensemble. To do so, simplicial depth fields were built for both the original ensemble and each of the selected subsets. Since simplicial depth provides a measurement of the

centrality of a point in a collection, comparing of two simplicial depth fields can be used to evaluate the representativeness of the subsets. These simplicial depth values are normalized to the range $[0, 1]$ and color-coded for display. Figure 4.7a shows a simplicial depth field interpolated from the original ensemble, Figure 4.7b from the OLS subset, and Figuer 4.7c from the WSE subset. To enhance the visual comparison, the grayscale value of the OLS and WSE images were subtracted from that of the original ensemble image, the differences being displayed in Figure 4.7d and Figure 4.7e. Both of those subtractions result in a variation of simplicial depth values between $[-0.05, +0.05]$ out of a possible range of $[-1, +1]$, meaning the maximum variation is only 5% of the range of the simplicial depth values. These results indicate that both of these subsets are highly representative of the original sample set, and the choice of the selection algorithm can be left to the users based on their visualization requirements.




(a) Original ensemble (b) OLS (c) WSE

(d) Original - OLS (e) Original - WSE

Figure 4.7: Comparison of simplicial depth fields of the original ensemble with subsets selected via the OLS and WSE algorithms

Because the subsets selected by this approach are highly representative of the original ensemble, they can be used to smoothly interpolate spatio-temporal information represented by the original ensemble

to generate summary visualizations. Interpolating the simplicial depth field of a selected subset can produce a visualization showing the geospatial spread of the likelihood of a hurricane strike at a particular point in time using concentric ellipsoidal confidence intervals. Chapter 3 developed confidence interval visualizations by fitting ellipses to the simplicial depth field computed from the full ensemble, the top row of Figure 4.8 showing a time sequence produced utilizing this approach. To give this method coherency over time, it requires the use of a non-linear stabilizing filter to reduce temporal instability in the ellipse orientation. Such a filter potentially introduces arbitrary errors into the visualization. The approach developed by the work reported here using the OLS selection algorithm produces similar visualizations directly from the interpolated simplicial depth field without requiring an ellipse fitting, meaning the results are not only faster to compute but remain stable without the use of a filter when viewed as a time sequence. The bottom row of Figure 4.8 shows such a time sequence produced utilizing our algorithm.

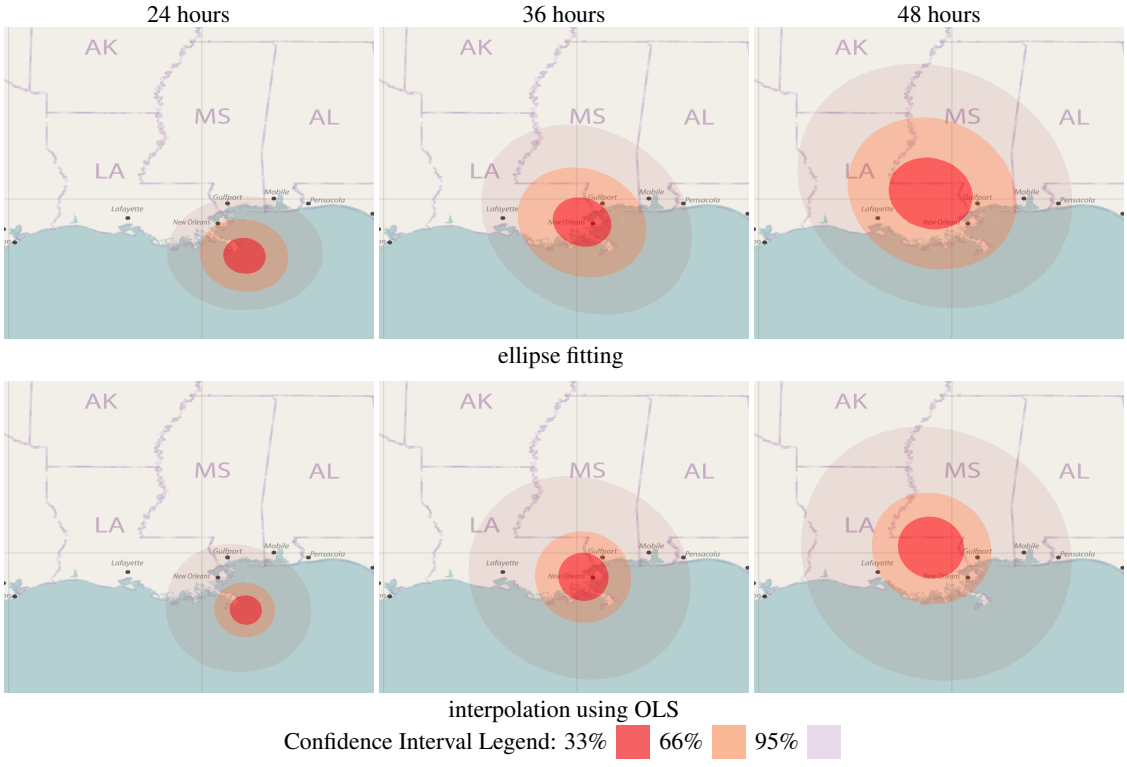
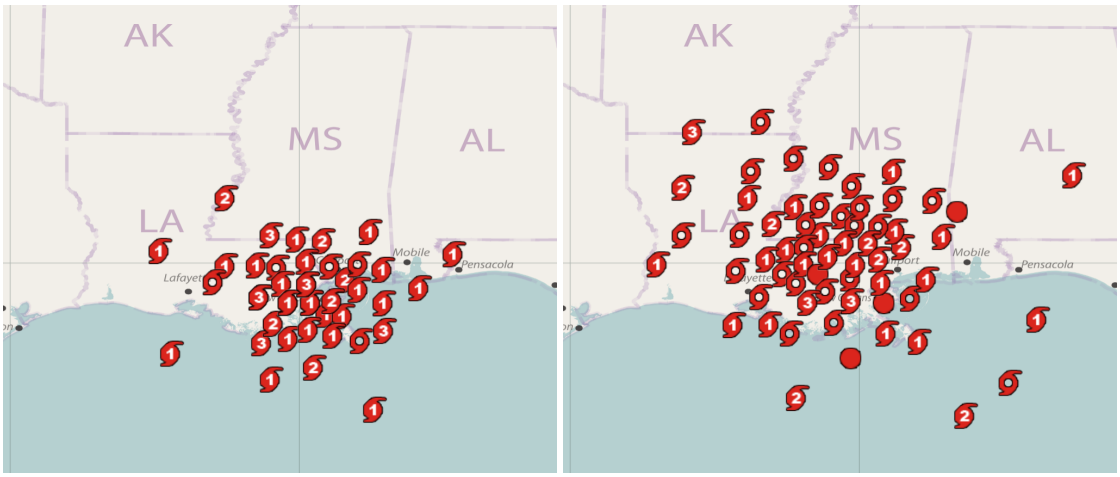



Figure 4.8: Simplicial depth field displays over time, comparing results generated in Chapter 3 with interpolation from an optimal subset

The main goal of this work was to produce uncluttered, representatively spaced, direct ensemble displays. One compelling attribute of this type of display is that glyphs carrying information in addition to position can be placed at the location of each sample with less concern for occlusion. To demonstrate that

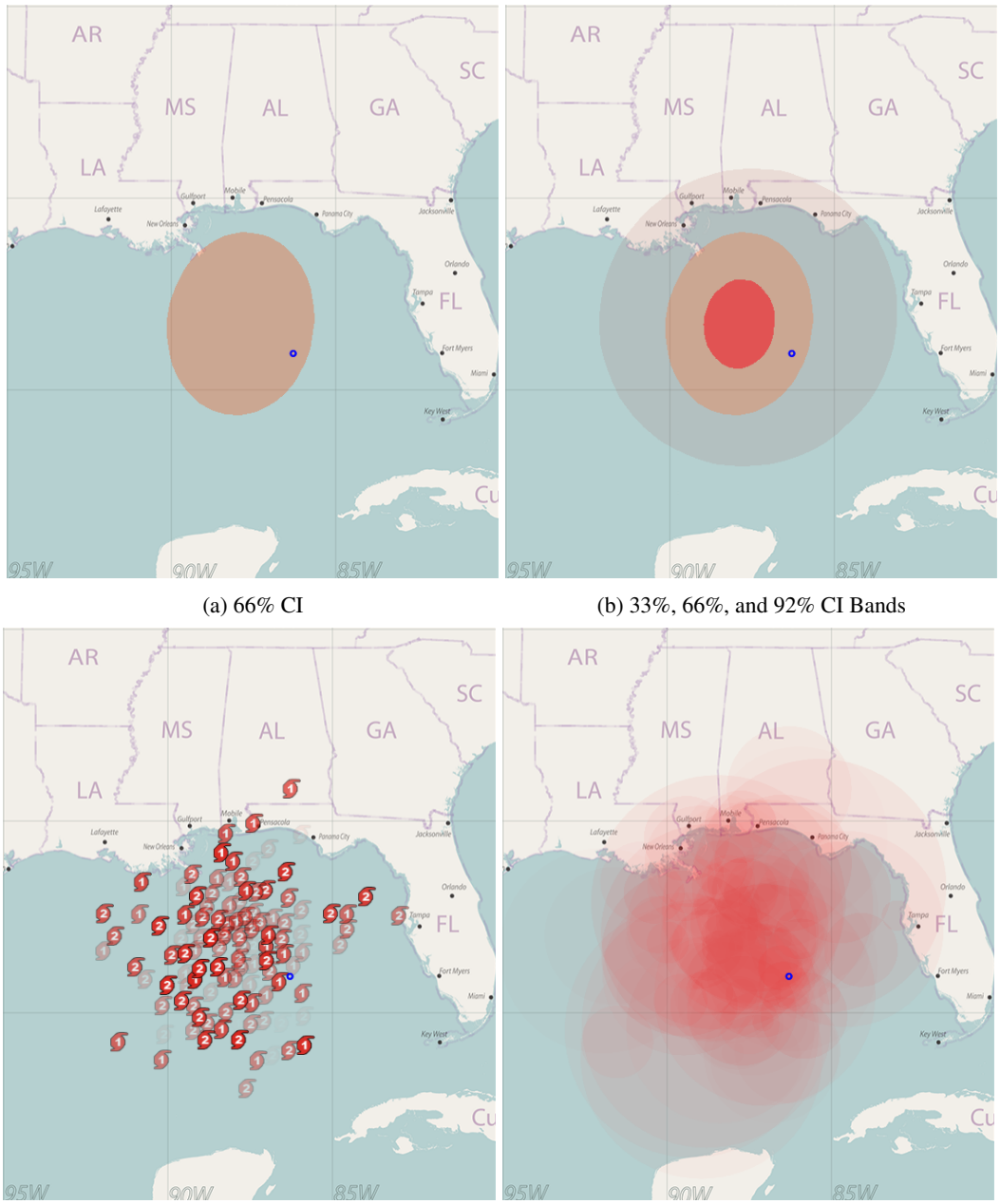
this technique has the potential to support this type of visualization, a direct ensemble display was created by placing glyphs indicating predicted storm intensity at each of the predicted locations. Figure 4.9, which shows two such displays at 36 and 48 hours into the prediction, indicate the most likely storm positions as well as that outliers are more likely as time progresses into the prediction. Since the goal was to clearly show outliers, the WSE algorithm was applied to obtain the subsets used in this particular case. However, according to meteorologists at the NHC, this visualization does not have enough coverage of the area likely to experience hurricane force winds. A second concern is that individual elements of the ensemble are highly salient, potentially misleading viewers into paying too much attention to individual glyphs rather than the overall distribution.




(a) 36 hours (b) 48 hours

Figure 4.9: Glyph displays of the prediction over time using ensemble subsets from the WSE algorithm

To address these concerns, an animated visualization, one continuously adding new ensemble members to the display while slowly fading out glyphs that have been on the screen for a while was created. Over time this allows many more ensemble members to be displayed without creating clutter and, since they are always fading away, also deemphasizes any particular glyph. Within this visualization, areas with a high concentration of samples tend to be more opaque than those with low concentrations, meaning more opacity indicates a higher level of certainty. This visualization was implemented by initially randomly specifying opacities for all of the samples, which were divided into two groups, one with non-zero opacity that is displayed, and one with zero opacity that is not displayed. Each time the opacity of a displayed sample decreases to zero, a new sample is selected from the zero opacity group, the simplicial depth value of which is closest to that sample, and displayed at its maximum opacity. This ensures that the distribution at each time frame




(a) 66% CI

(b) 33%, 66%, and 92% CI Bands

(c) Icons

(d) Sizes (Splats)

Figure 4.10: Four visualizations studied in the cognitive experiment. The blue dot indicates the position of an oil rig platform.

is as close to the distribution of the full ensemble as possible. Figure 4.10c shows a snapshot of such a visualization.

Another comment from the NHC meteorologists was that even though the glyphs depict a set of predicted positions of the center of the storm annotated by storm strength, they do not show the extent of the area potentially affected by hurricane force winds. To address this concern, a visualization using circles centered at the selected samples, the radii of which correspond to predicted distances from the center at which 50 kn winds are predicted, was created. The same pipeline was used for drawing strength glyphs to create a display that dynamically updates. A snapshot of this type of visualization is shown in Figure 4.10d.

## 4.4 Experimental Evaluation

To evaluate the effectiveness of the visualizations developed here in helping users estimate hurricane risk, we conducted a cognitive experiment in collaboration with researchers from the University of Utah and the University of California, Santa Barbara comparing each of these three new visualizations with a visualization showing only the 66% confidence interval for the expected storm center. The 66% CI was selected as the standard for comparison because it is most similar to current practice in communicating hurricane forecasts (i.e. the NHC uncertainty cone).

This cognitive experiment adopted historical NHC predictions of five US Gulf Coast hurricanes, three of which were also translated slightly to ensure that the predicted storm center was offshore at 48 hours from the time of the forecast. Four visualizations taken at 48 hours were generated for each storm. Two of these visualizations were static summaries of the hurricane locations, one is the 66% confidence interval (66 CI) (Figure 4.10a) and second the 33%, 66%, and 92% concentric confidence intervals (CI Bands) (Figure 4.10b). The remaining two were animated displays of ensembles of points, both representing the storm location, with one using glyphs to show the storm strength (Icons) (Figure 4.10c) and the other representing the storm size (Splats) (Figure 4.10d).

### 4.4.1 Experimental Methodology

This study involved 133 participants from the University of California, Santa Barbara and the University of Utah who were randomly assigned to a visualization type and provided with its description and instructions for the task. Adopting the experiment conducted by Ruginski et al. [RBP+16], participants were shown one of the forecast display techniques that included a superimposed blue dot indicating the location of

an offshore oil platform (see Figure 4.10). Displays developed using E-Prime 2.0.10.353 software [SEZ12] were presented on either a Dell U2412MB monitor or an Asus VG248 monitor in sRGB color mode with pixel resolutions of $1920 \times 1080$. The offshore oil platform was located at one of eight locations defined relative to the center of the forecast. After being given a description of the oil platform, the participants were instructed to estimate the level of damage it would incur based on the likelihood of the storm affecting it and the strength of the storm in the affected region. Below the display was a Likert scale ranging from 1 (no damage) to 9 (severe damage) for the participants to use to enter their responses on a keyboard. This scale was chosen to provide participants with a means of hurricane interpretation that would simulate a real-world decision task. Each completed 40 trials in which the locations of both the platform and hurricane were randomly presented for their visualization type only. Finally, they were asked a series of seven true/false statements to assess their understanding and misconceptions of the displays.

A multilevel model (MLM) was fitted to the damage judgment data using Hierarchical Linear Modeling (HLM 7.0) software and restricted maximum likelihood estimation procedures [RABCdT11, RB02]. MLM is a generalized form of linear regression used to analyze variance in experimental outcomes on both individual (within-participants) and group (between-participants) levels. An MLM was appropriate for these data and testing these hypotheses for two reasons: 1) it allows for the inclusion of interactions between continuous variables (in this case, distance) and categorical predictors (in this case, the type of visualization); 2) it uses estimation procedures appropriate for partitioning variance and error structures in mixed and nested designs (repeated measures nested within individuals in this case).

The multilevel model used here:

$$\begin{aligned}\text{Damage}_{ij} = {} & \gamma_{00} + \gamma_{01} \times \text{Icons}_j + \gamma_{02} \times \text{Splats}_j + \\ & \gamma_{03} \times \text{CIBands}_j + \gamma_{10} \times \text{Distance}_{ij} + \\ & \gamma_{11} \times \text{Icons}_j \times \text{Distance}_{ij} + \gamma_{12} \times \text{Splats}_j \times \text{Distance}_{ij} + \\ & \gamma_{13} \times \text{CIBands}_j \times \text{Distance}_{ij} + u_{0j} + r_{ij},\end{aligned}$$

where $i$ represents trials and $j$ individuals, $\gamma$ terms are the regression coefficients, the error term $r_{ij}$ represents the variance in the outcome variable on a per trial basis and $u_{0j}$ on a per person basis [CCWA13].

Damage rating, although an ordinal variable by definition, was treated as a continuous one in the model because it contained more than five response categories [BS11]. For the distance variable, the absolute value of the oil rig distances, regardless of which side of the hurricane forecast they were on, was analyzed,

Table 4.1: Average damage judgments made at the center of the forecasts. Main effect (intercept) references 66% CI visualization. The table summarizes the comparison of each visualization to the 66% CI.

| Fixed Effect | Coeff. | Std Err | $t$-ratio | DOF | $p$-value | 95% CI |
|---|---|---|---|---|---|---|
| Intercept 1, $\beta_0$ | | | | | | |
| Intercept 2, $\gamma_{00}$ | 9.129 | 0.177 | 51.68 | 119 | <0.001 | (8,78,9.48) |
| Icons, $\gamma_{01}$ | -2.525 | 0.371 | -6.79 | 119 | <0.001 | (-3.26, -1.79) |
| Splats, $\gamma_{02}$ | 0.159 | 0.221 | 0.72 | 119 | 0.473 | (-0.27, 0.59) |
| CI Bands, $\gamma_{03}$ | -1.433 | 0.353 | -4.06 | 119 | <0.001 | (-2.12, -0.74) |

as none of our hypotheses related to whether they were located on a particular side. This distance was divided by 10 prior to analysis so that the estimated model coefficient would correspond to 10 km changes (rather than a 1 km change). The type of visualization was dummy-coded such that this 66% CI was the reference group, allowing for comparison of each visualization to the 66% CI. The analysis was subsequently collapsed over the five hurricane forecasts.

The mixed two-level regression model tested whether the effect of distance from the center of forecasts (level 1) varied as a function of visualization (level 2). It was hypothesized that the damage judgments made by the participants who viewered the CI Band, Icon, and Splat visualizations would change more gradually as a result of distance compared to the 66% CI, demonstrating interpretations more in line with an uncertainty distribution.

### 4.4.2 Damage Judgment Results

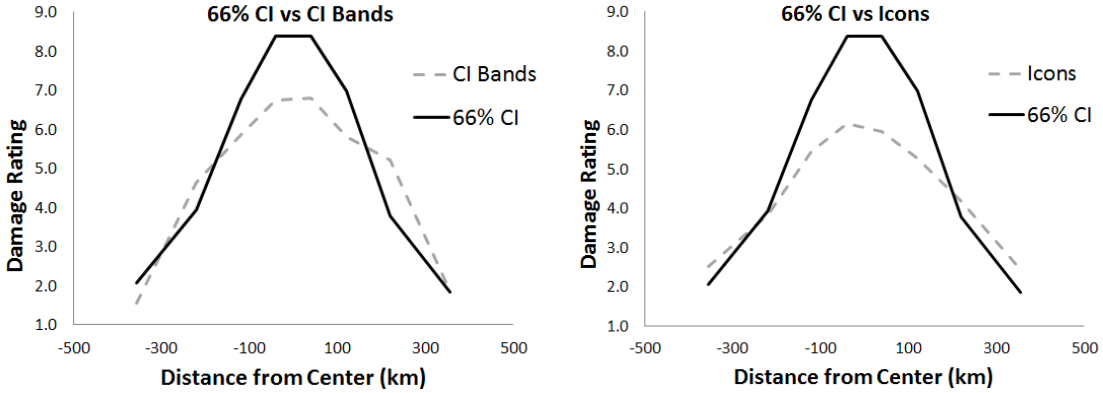
The analysis indicated that the average damage ratings were approximately the maximum possible (9.13) at the center of the 66% CI as seen in Table 4.1. This value is estimated above the scale since individuals never made a damage judgment exactly at the center of the hurricane. Compared to the 66% CI, damage judgments made using the CI Bands at the center of the hurricane forecast were on average 1.43 lower, and judgments made using the Icon were on average 2.53 lower. However, average damage ratings made at the center of the forecast using the Splat visualization were not meaningfully different from those made using the 66% CI.

Secondly, the analysis revealed a significant association between the distance and damage rating for the 66% CI visualization (0.21 decrease in damage judgment per 10 km on average as seen in Table 4.2), a result that was expected given the participants had little information aside from distance when making damage judgments using this visualization. More importantly, both the CI Bands visualization (0.157 decrease per 10 km, as seen in Figure 4.11a) and the Icon visualization (0.116 decrease per 10 km, as seen in Figure 4.11b)

Table 4.2: Effect of a 10 kilometer change in distance on damage ratings. Main effect (intercept) references 66% CI visualization. The table summarizes the comparison of each visualization to the 66% CI.

| Fixed Effect | Coeff | Std Err | $t$-ratio | DOF | $p$-value | 95% CI |
|---|---|---|---|---|---|---|
| Slope, $\beta_0$ | | | | | | |
| Intercept 2, $\gamma_{00}$ | -0.210 | 0.006 | -32.36 | 4793 | <0.001 | (-0.22,-0.19) |
| Icons, $\gamma_{01}$ | 0.094 | 0.015 | 6.31 | 4793 | <0.001 | (0.06, 0.12) |
| Splats, $\gamma_{02}$ | -0.003 | 0.009 | -0.36 | 4793 | 0.720 | (-0.02, 0.01) |
| CI Bands, $\gamma_{03}$ | 0.053 | 0.018 | 2.94 | 4793 | 0.003 | (0.02, 0.09) |

demonstrated a significantly weaker association between the distance and damage rating than the 66% CI. However, the relationship between distance-from-center and the damage rating for the Splats visualization did not differ significantly from the 66% CI visualization.




(a) 66% CI v.s CI Bands

(b) 66% CI v.s Icons

Figure 4.11: Mean damage ratings as a function of distance from the center of the storm. Each visualization is compared to the 66% CI.

### 4.4.3 Post-task Questionnaire Results

In the post-task survey, the participants were asked a series of seven true/false statements (given in Table 4.3), comparing the responses from each of the other 3 visualization conditions to those of the 66% CI condition using chi-squared tests of independence. Responses to Statement 1 indicated that participants in the CI Bands group, $\chi^2(1, N = 65) = 4.99, p = 0.03$, and participants in the Splats condition, $\chi^2(1, N = 62) = 10.09, p = 0.001$, were more likely to agree with the statement that the forecast shows a distribution of possible locations. Responses to Question 2, 3, and 4 did not differ significantly, indicating that across all conditions, most participants understood that the displays indicate 1) that the area being shown has a chance of being damaged, 2) where the center of the storm is likely to be, and 3) that the forecasters are uncertain of

the storm's location. More importantly, although only the Splats condition showed the size of the hurricane, there was no significant difference among conditions for Statement 5 that the forecast visualization indicates the size of the hurricane. Participants in the 66% CI in particular exhibited similar agreement rates with this statement with those in the Splats condition, indicating a strong misconception that the size of the glyphs indicates the size of the hurricane.

Table 4.3: Percentage of participants in agreement (out of total answering). Approximately 1/3 of the respondents were in each visualization category.

| Statement | 66% CI | CI Bands | Splats | Icons |
|---|---|---|---|---|
| 1) The forecast shows a distribution of possible hurricane locations | 63.6% | 87.5% | 96.6% | 82.8% |
| 2) The forecast shows the area that has a chance of being damaged | 78.8% | 87.5% | 93.1% | 92.9% |
| 3) The forecast shows where the center of the hurricane has a chance of being | 87.9% | 87.5% | 93.1% | 92.9% |
| 4) The forecast visualization indicates that the forecasters are not certain about the location | 66.7% | 68.8% | 62.1% | 44.8% |
| 5) The forecast visualization indicates the size of the hurricane | 57.6% | 40.6% | 58.6% | 37.9% |
| 6) The forecast visualization indicates the intensity of the hurricane | 9.1 % | 43.8% | 71.4% | 96.6% |
| 7) The forecast visualization indicates the passing of time | 0.0 % | 6.3 % | 41.4% | 41.4% |

Additionally, this study did not find strong misconceptions that the size of the glyph in the 66% CI condition indicates intensity (see responses to Statement 6). More importantly, most participants in the Icon display condition correctly indicated that this display showed intensity, significantly different from the 66% CI condition, $\chi^2(1,N=62)=47.23, p<0.001$. However, the CI Bands condition, $\chi^2(1,N=65)=10.11, p=0.001$, and the Splats condition, $\chi^2(1,N=62)=25.06, p<0.001$, were also more likely than the 66% confidence interval group to agree with the statement that the displays showed the intensity of the hurricane although neither of these displays actually does so. Finally, although none of these displays shows the passage of time, participants in the two animation conditions (Splats, $\chi^2(1,N=62)=16.93, p<.001$; Icon, $\chi^2(1,N=62)=16.93, p<0.001$) were more likely to agree with Statement 7 compared to those in the 66% CI condition, a result suggesting a possible misconception that change of time in a display indicates change over time in the forecast.

### 4.4.4 Experiment Summary

As this analysis suggests, the CI Bands and Icon visualizations were interpreted as changing more gradually as a function of distance than the 66% CI visualization, suggesting that the former two were interpreted more like an uncertainty distribution than the latter, a result similar to the effect seen with the path ensemble displays used in the Ruginksi et al.'s experiment [RBP+16]. However, the Splats visualization was interpreted similarly to the 66% CI, both in terms of average overall damage rating provided across all trials and the change in damage ratings that occurred as a function of distance-from-center.

Even though the subset selection method developed here works well for point ensembles, it cannot be directly applied to ensembles of tracks. In the beginning of this chapter, experimental evidence that path ensembles have the potential to be used for making an effective visualization was discussed. Therefore, the next goal of the research introduced here is to develop a method that reduces the number of paths in an ensemble by making a selection that preserves the statistical properties of the ensemble and, thus, to produce visualizations of ensembles that improve upon the current visually cluttered displays.

# Chapter 5

# Visualizing Hurricane Predictions with Uncertainty by Representatively Sampling Ensembles of Hurricane Forecast Paths

As discussed in the previous chapter, directly displaying ensembles of the predicted tracks of hurricanes has the potential to minimize the confusion concerning the spatial attributes of a prediction with its uncertainty [CHL13, RBP+16, CG15]. However, visualizing all members of such ensembles typically leads to overplotting, thus resulting in viewers having difficulty interpreting the information depicted. More importantly, this limitation prevents visualization designers from incorporating such conventional approaches as modifying colors or adding glyphs to portray the parameters of the storm. For instance, tracks are usually visually encoded using polylines, the colors of which can indicate possible variations of such meaningful information as the intensity and potential damage ratings of the individual tracks. In addition, well-designed glyphs can be attached to time-sampled locations along each track to indicate more spatial-temporal information. By doing so, the viewers are provided with both an overall view and the time-specific details of a forecast in a single display. However, these geometrical manipulations are potentially unavailable to the viewers in cluttered visualizations.

The research reported here addresses these limitations by displaying a small collection of representative hurricane-predicted tracks obtained by reconnecting time-specific hurricane-predicted locations included in the original ensemble of forecast tracks. Similar to the method discussed in Chapter 4, such a small collection of tracks is considered to be representative if the following two requirements are satisfied:

- Its members accurately preserve the original distributions of available storm characteristics including the location (latitude and longitude), the storm intensity, and the storm size. This requirement is crucial because a visualization portraying the forecast incorrectly is not only meaningless but also dangerous for guiding the emergency decision-making needed during hurricanes.
- Its members should maintain an appropriate spatial layout, meaning intersections between tracks should be minimized.

Primarily, the technique reported here for achieving these goals begins with extracting a median track of an ensemble by calculating its non-parametric statistical spatial characteristic, partitioning the ensemble into two equivalent groups, then constructing the representative set by recursively computing subsequent median tracks from successive groups. Even though these median tracks are not the original forecast tracks included in the ensemble, using them to convey the hurricane forecast does not decrease the accuracy of the visualization developed here. Recall that these forecast track ensembles are generated by Monte-Carlo sampling statistical models and thus, using individual ensemble members does not provide either significant statistical information of the hurricane forecasts or the realistic physical behaviors of the hurricanes. The most important information of such an ensemble is the overall statistical distribution portrayed by all of the ensemble members. The extracted median tracks preserve this distribution at individual points in time by including the most representative predicted locations, i.e. the median locations, and inherently, preserve the spatial distribution over the complete forecasting time. However, any method using tracks parameterized by time cannot preserve the hurricane speed uncertainty distribution. This is a limitation of the method developed here.

In addition, to include other variables in the ensemble, the RBF interpolation systems with adaptive kernel spreads introduced in Chapter 3 are built for each of them, allowing for assigning of the values for these variables to time-sampled locations along tracks in the subset. Subsequently, a visualization portraying NHC forecast ensembles by showing representative tracks computed using this algorithm is designed. The predicted storm intensities are visually encoded using a carefully designed color scheme, while the predicted storm sizes are represented by circles, the radii of which are determined by the spatial spread of the areas

predicted to experience the hurricane.

Finally, a cognitive study investigating viewers' interpretations of this visualization by examining their estimations of potential damages to simulated oil rig platforms was conducted. The preliminary experimental results revealed that this visualization was interpreted by the participants more as an uncertainty distribution than misunderstanding that the hurricane was predicted to be strongest in the spatial center of the forecast.

## 5.1 Methodology

The following describes the details of the method developed for computing a representative subset from an ensemble of forecast tracks. Mathematically, an ensemble $\mathbf{E}$ of $m$ tracks is defined as $\mathbf{E} = \{\mathbf{D}_1, ..., \mathbf{D}_m\}$ and a member track $\mathbf{D}_i$ $(1 \leq i \leq m)$ is described by a set of data points $\mathbf{D}_i = \{\mathbf{d}_{i,0}, ..., \mathbf{d}_{i,T}\}$, where

- $T$ defines the number of samples along a track, e.g. in the NHC ensemble sampling 120-hours prediction at every hour, $T$ equals 121;
- a data point $\mathbf{d}_{i,j} \in \mathbf{D}_i$ is a collection of attributes: $\mathbf{d}_{i,j} = \{t_{i,j}, p_{i,j}, A^1_{i,j}, ..., A^S_{i,j}\}$;
- $t_{i,j}$ indicates the time of this point;
- $p_{i,j}$ denotes its spatial coordinates, e.g. latitude and longitude;
- $\{A^1_{i,j}, ..., A^S_{i,j}\}$ represents numerical values of $S$ quantitative attributes of the data point, e.g. in a hurricane forecast ensemble including storm size and intensity, $S$ equals 2.

The steps of the framework introduced here for computing a subset $\tilde{\mathbf{E}}$ of $n < m$ tracks from ensemble $\mathbf{E}$ is summarized below, followed by a visual depiction shown in Figure 5.1:

I Extract a median track $\tilde{\mathbf{d}}$ from $\mathbf{E}$ as shown by the white line in Figure 5.1(a).

The fundamental concept of this algorithm for reducing the number of tracks is to depict the entire ensemble using its most significant representation, i.e. a *median track* formed by connecting the median points at each time step, as described below in Section 5.1.1.

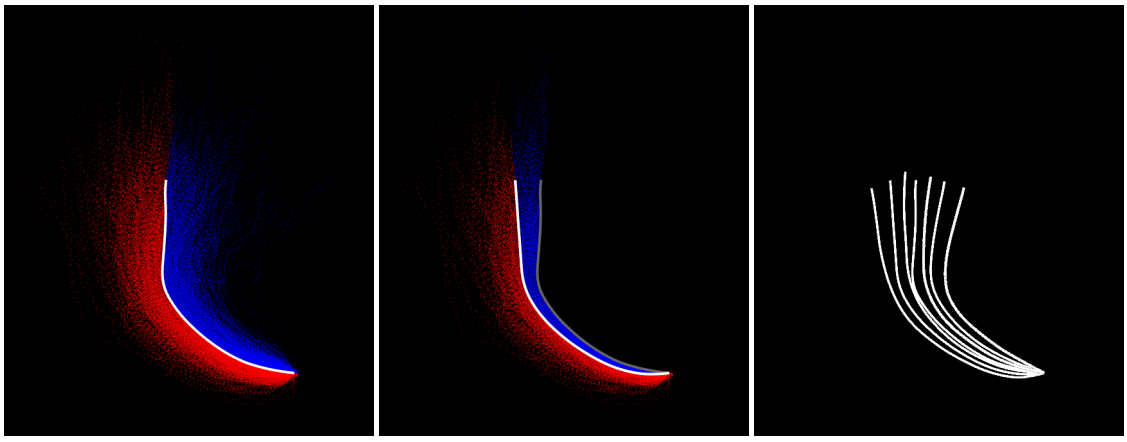
II Equally divide $\mathbf{E}$ into two new smaller ensembles, $\mathbf{E}_L$ and $\mathbf{E}_R$, as shown by the red and blue points in Figure 5.1(a).

The median track enables a determination of the left and right regions of the spatial spread covered by the ensemble, thus enabling an equal partition of the spatial locations as described below in 5.1.2.

III Recursively construct median tracks from $\mathbf{E}_L$ (Figure 5.1(b)) and $\mathbf{E}_R$ until achieving $n$ tracks (Figure 5.1(c)).

By recursively applying the previous two steps on both the left and right partitions, this algorithm consecutively extracts more paths in a descending order of their representativeness of the original ensemble, the resulting collection of $n$ median tracks determines the final subset $\tilde{\mathbf{E}}$, as described below in 5.1.3.

IV Assign appropriate ensemble attributes to the data samples along the median tracks in $\tilde{\mathbf{E}}$ by using RBF interpolations over the spatial spread of the original ensemble $\mathbf{E}$ as described below in Section 5.1.4. This approach preserves the original statistical distributions of these ensemble attributes.



(a) (b) (c)

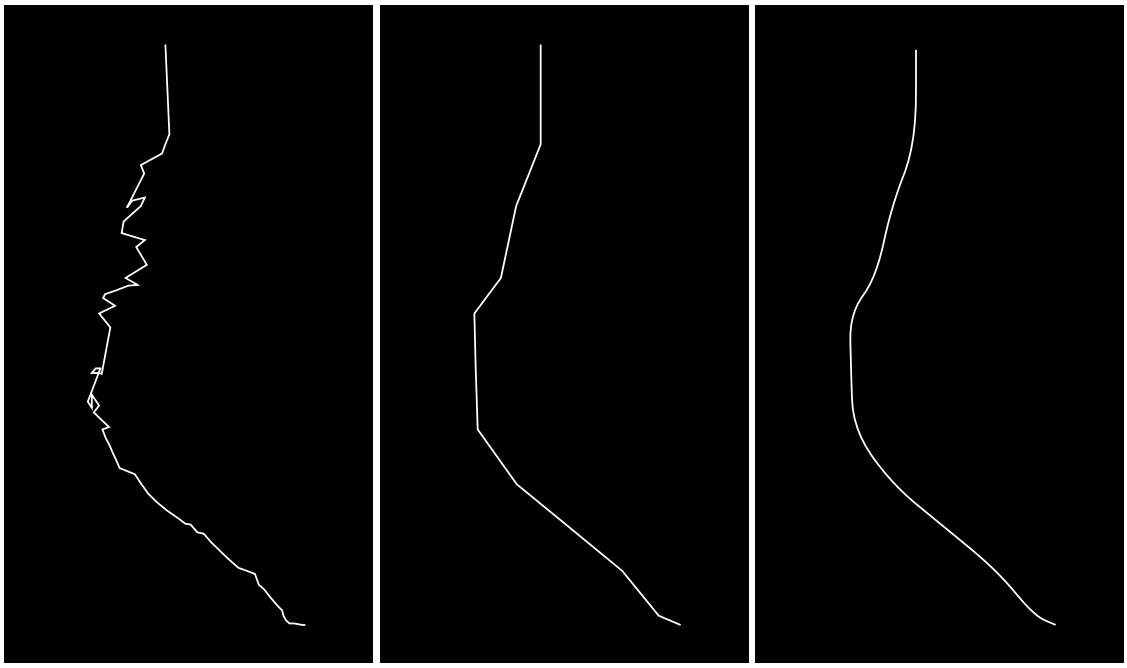
Figure 5.1: A demonstration of the proposed recursive sampling process for extracting median tracks from an ensemble: (a) exemplifies the first level of recursion, where the colored points represent all spatial locations included in an ensemble, the white line indicates the median track extracted from this ensemble and the red and blue are used for labeling points belonging to the left and right groups separated by the median track; (b) shows a second-level recursion on the left group obtained from the first-level recursion; and (c) demonstrates a collection of 7 median tracks computed by running 3 levels of recursion.

### 5.1.1 Constructing a Median Track from an Ensemble

The algorithm reported here begins by extracting collections of spatial points at individual points in time from ensemble $\mathbf{E}$, e.g. the points set at time $j$ is $\{p_{1,j}, ..., p_{m,j}\}$. Then the median of this set is computed

by selecting the deepest point with the largest simplicial depth value. Since simplicial depth provides a center-outward ordering of a collection of points, the deepest point indicates the spatial center of these points, thereby most significantly representing the entire collection.

Subsequently, a median track of the ensemble is constructed by connecting individual median points over time. Figure 5.2(a) exemplifies a raw median path extracted from an NHC hurricane forecast track ensemble; as this figure shows, this path is highly irregular with self-intersections and zigzags, therefore it is difficult to determine its left and right sides. However, this determination is essential for the following recursive stages of the sampling algorithm. In addition, these irregularities have no crucial statistical significance as they are high frequency artifacts due to the high resolution of simulation time used to generate the NHC ensemble; as a result, this raw median path needs to be smoothed by removing these irregularities.



(a) (b) (c)

Figure 5.2: Process of smoothing a median track: (a) shows a raw median path extracted from an ensemble which has irregularities due to high frequency artifacts of simulation, (b) shows a simplified version of the raw median track using the Lang algorithm, and (c) shows the final median track generated by b-spline curve fitting the simplified track.

The first step in smoothing a path is to process the self-intersections, which in the context of hurricane forecasting can occur over a short period of time, e.g. hours, or a long period encompassing days. The former should be eliminated because an actual hurricane can rarely exhibit such a large variation in bearing

and speed that it crosses its previous track in only a few hours. However, those within long periods of time should be preserved because this situation can often occur in long-term predictions. (For instance, the prediction of Hurricane Matthew at 5PM, October 6, 2016, forecast that it would reverse and hit Florida twice within the next five days.) To implement this process, a loop detection over a median track with a time window is executed. Typically, if a loop is detected and the time interval between its starting and ending points falls within the window, it will be eliminated by directly reconnecting the starting and ending points, ignoring all intermediate ones. If this criterion is not met, the loop will be kept.

Second, the zigzags of a track are removed by using polyline simplification. The work reported here uses the Lang algorithm [Lan69] due to its advantages of maintaining the original points, avoiding new self-loops, and retaining positional accuracy of the original line [SC06]. Specifically, this Lang algorithm begins by defining a search region including a fixed number of consecutive points of the original line and then constructing a tentative simplified segment by connecting the first and last points of this region. Subsequently, the perpendicular distances from these points to the tentative simplified segment are computed. If a distance is greater than a predefined threshold, the search region is resized by excluding the last point, and the corresponding point-to-segment perpendicular distances are computed again. This process repeats until no distance exceeds the threshold. Then, all intermediate points of this search region excluding the first and last ones are eliminated, and a new search region is formed starting with the last point of the former search region. The algorithm continues until reaching the last point of the original line. Figure 5.2(b) shows the simplified version of the raw median track.

However, this simplification results in losing forecasting time conveyed by these eliminated points. To address this issue, a B-spline curve [Sch46] is fitted from the simplified path. A B-spline curve consists of a sequence of polynomial curve segments, each of which is usually evaluated using a de Boor algorithm [DB72]. More specifically, given the degree $n$ of each curve segment, the knot vector $U = \{u_0, ..., u_K\}$ ($u_i \leq u_{i+1}$) and a sequence of control points $d_0, ..., d_L$, where $L = K - n + 1$, the piecewise polynomial $s$ in interval $[u_{n-1}, u_L]$ of a B-spline curve is defined as

$$s(u) = \sum_{i=0}^{L} d_i N_{i,n}(u), \tag{5.1}$$

where the B-spline basis function $N_{j,n}(u)$ is defined by the recursion formula:

$$N_{j,0}(u) = \begin{cases} 1 & \text{if } u_{j-1} \leq u < u_j, \\ 0 & \text{otherwise.} \end{cases}$$

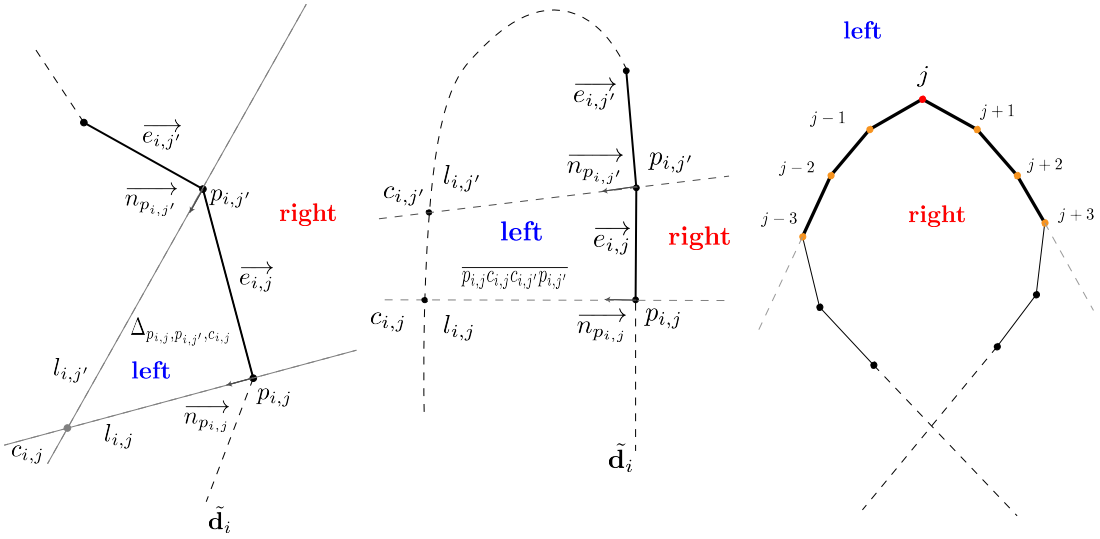
$$N_{j,n}(u) = \frac{u - u_{j-1}}{u_{j+n-1} - u_{j-1}} N_{j,n-1}(u) + \frac{u_{j+n} - u}{u_{j+n} - u_j} N_{j+1,n-1}(u).$$

In this work, the knot vector is determined by the time information attached to these locations. To evaluate the actual curve for visualization purposes, the B-spline parametric variable $u$ is specified as instances over the entire time interval of the forecast. For the NHC ensembles of 120-hour predictions, $u$ belongs to $[0, 120]$. A final version of the median track exemplified here is shown in Figure 5.2(c).

### 5.1.2 Partitioning an Ensemble

The next step in the sampling approach is to divide $\mathbf{E}$ into two equal smaller ensembles. To do so, normal directions at individual spatial points along the median track need to be defined. As shown in Figure 5.3(a) and (b), given two adjacent time points $j$ and $j'$ ($j' = j + 1$) on median track $\tilde{\mathbf{d}}_i$ and their corresponding spatial locations along the track, $p_{i,j}$ and $p_{i,j'}$, a directed line segment $\overrightarrow{e_{i,j}}$ from $p_{i,j}$ to $p_{i,j'}$ is constructed, and its normal direction $\overrightarrow{n_{e_{i,j}}}$ is computed. Consequently, the normal direction of a spatial location is defined as the normal direction of the directed line segment starting from it. For instance, the normal direction $\overrightarrow{n_{p_{i,j}}}$ of $p_{i,j}$ is identical to $\overrightarrow{n_{e_{i,j}}}$, while the normal direction $\overrightarrow{n_{p_{i,j'}}}$ of $p_{i,j'}$ is identical to the normal direction of the next succeeding segment $\overrightarrow{e_{i,j'}}$ starting from $p_{i,j'}$. The normal direction $\overrightarrow{n_{i,T}}$ of the last point of the track is identical to the normal direction at $p_{i,T-1}$.

Once the normal directions are defined, the left and right sides of individual directed segments can be determined. First, the median track is extended infinitely along its starting and ending directions. Then for each segment $\overrightarrow{e_{i,j}}$, two infinite lines $l_{i,j}$ and $l_{i,j'}$ starting from $p_{i,j}$ and $p_{i,j'}$ are derived along their normal directions $\overrightarrow{n_{p_{i,j}}}$ and $\overrightarrow{n_{p_{i,j'}}}$. If $l_{i,j}$ and $l_{i,j'}$ intersect at $c_{i,j}$, the left region of this segment is defined as the area enclosed by the triangle $\Delta_{p_{i,j},p_{i,j'},c_{i,j}}$ shown in Figure 5.3(a). If $l_{i,j}$ and $l_{i,j'}$ intersect with the extended median track at $c_{i,j}$ and $c_{i,j'}$, the left region of this segment is defined as the area enclosed by the polygon $\overline{p_{i,j}c_{i,j}c_{i,j'}p_{i,j'}}$ shown in Figure 5.3(b). If neither of these two situations is satisfied, the left region is defined as the area enclosed by $\overrightarrow{e_{i,j}}$, $l_{i,j}$, and $l_{i,j'}$. Similarly, the right region of $\overrightarrow{e_{i,j}}$ is defined by constructing $l_{i,j}$ and $l_{i,j'}$ following the negative normal directions. The left/right region of the median track is then defined as the




(a) (b) (c)

Figure 5.3: Determination of left and right regions of a directed line segment of a median path: (a) exemplifies the case that $l_{i,j}$ and $l_{i,j'}$ intersect each other, while (b) exemplifies the case that $l_{i,j}$ and $l_i, j'$ intersect with $\tilde{\mathbf{d}}_i$, and (c) demonstrates the idea of time window we used to address the difficulty of determining left and right regions when a median path has a long term loop. Details are described in the text.

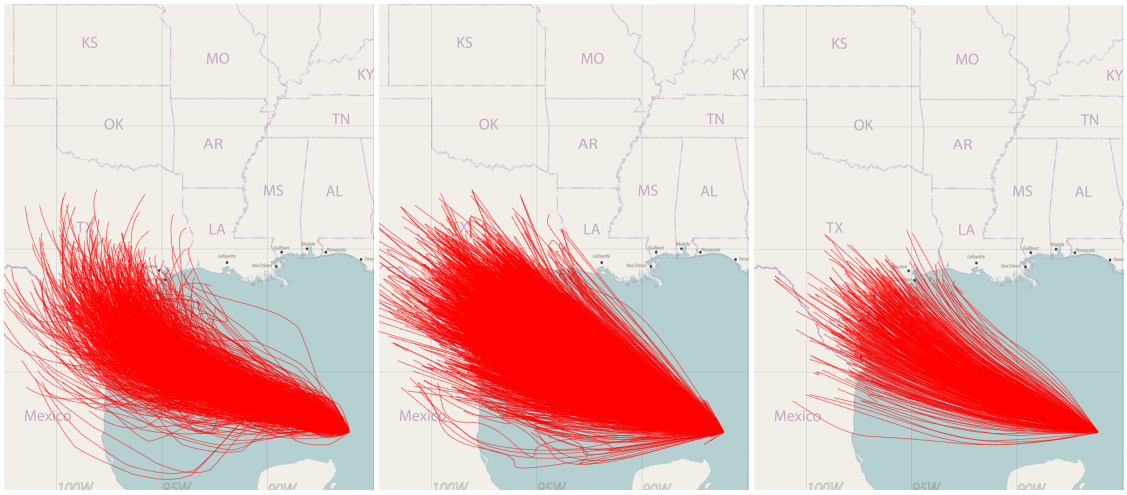
union of corresponding left/right regions of all directed line segments along the track.

However, since the process of smoothing a median track preserves long-term spatial topology, the processed track potentially includes loops indicating the actual long-term reverse of a hurricane, resulting in the problem that left regions of some directed line segments overlap with right regions of others, leading to inherent ambiguity in determining the left and right of the track. This issue is addressed by integrating the time window introduced in 5.1.1. As demonstrated in Figure 5.3(c), for a sample point at time $j$, shown by the red point, is centered in a time window indicated by the orange points. The determination of the left and right regions of the path at time $j$ is constrained to consider only the left and right regions of directed line segments within this time window, outlined by the bold black lines. Subsequently, individual time-specific locations of $\mathbf{E}$ can be grouped into the left/right partition based on their corresponding time-window-constrained left/right area of the median track.

### 5.1.3 Constructing a Representative Set of Tracks

To acquire $n$ median tracks, the previous two steps are recursively executed for the original full ensemble $\mathbf{E}$ and its following subdivisions $\mathbf{E}_L^k$ and $\mathbf{E}_R^k$, where $k$ denotes the level of recursion and $k \geq 1$. Since the number of median tracks generated at the $k$th level of recursion is $2^k - 1$, the process is terminated when $2^k - 1 \geq n$. Because any specific median track divides $\mathbf{E}$ or any succeeding portion of it into two equal halves and it is the most representative of its portion spatially and statistically, a collection of these median tracks represents $100\% \times \sum_2^k \frac{1}{2^{k-1}}$ of the statistical spatial distribution conveyed by the full ensemble. For instance, 7 tracks generated from 3 levels of recursion cover approximately 75% of the original spatial distribution, and when $n = m$, the subset covers the original full ensemble.

However, this strategy potentially produces an increasing number of invalid median tracks as few points exist in partitions when the recursion hierarchy increases. Such tracks are constructed from a small number of points, inaccurately depicting the behavior of the hurricane. Figure 5.4(a) shows an NHC forecast track ensemble, while (b) shows 1,000 median tracks extracted from it using the approach developed here; as these two illustrations show, Figure (b) conveys a wider spatial spread than that implied by Figure (a) because some of the tracks resemble straight lines, a result of the smoothing operations introduced in the previous section computing an approximation of a median track, inducing a spatial offset between the raw and processed tracks. This issue is demonstrated in Figure 5.5, where the gray line represents a raw median track and the smoothed result is shown by the yellow line. In addition, when the algorithm reaches lower levels of recursion, the extracted partitions contain fewer samples such that at some points in time no samples




(a) (b) (c)

Figure 5.4: A demonstration addressing the problem of straight median tracks: (a) shows an NHC forecast track ensemble, (b) shows 1,000 unfiltered median tracks extracted from (a), and (c) shows 880 filtered tracks from (b).

can be assigned to either the left or right group of the smoothed median track as exemplified by the red points in Figure 5.5. Succeeding recursions, thus, will generate tracks that are empty at these points in time, making the track nearly identical to a straight line when there are many such time points.

The method developed here filters out these invalid tracks by measuring the number of valid points conveyed. More specifically, a median track is eliminated if it contains no legitimate spatial points at $fT$ time steps, where $f$ is a user-defined fraction set to 0.6 for the NHC forecast ensembles. A filtered collection of 880 tracks is shown in Figure 5.4(c) where the primary trend of the forecast depicted in Figure 5.4(a) has been maintained.

### 5.1.4 Incorporating Ancillary Ensemble Variables

The ultimate goal of the approach developed here is to incorporate accurately all available ensemble variables in a smaller subset of tracks. This goal is achieved by applying the RBF interpolation technique introduced in Chapter 3. For instance, to include predicted storm intensity at a specific point in time $j$, this approach computes a smooth, continuous intensity field by establishing an RBF interpolation system, the input samples of which are spatial locations at time point $j$; the known values of the input samples are their associated predicted intensities. To avoid the over-fitting problem due to the large number of spatial locations, a small, representative subset is extracted from this collection of locations using the Orthogonal

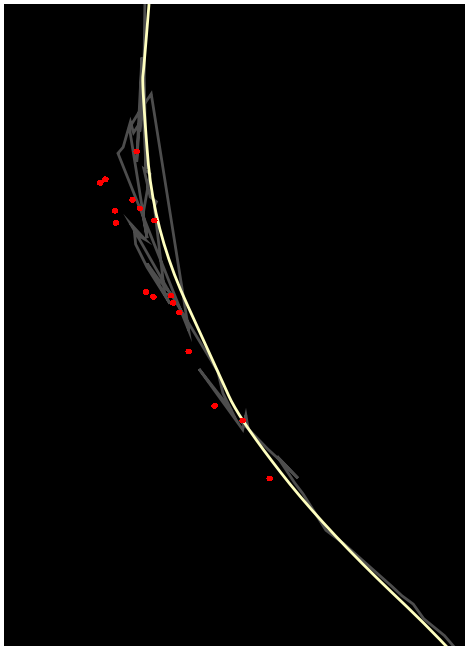


Figure 5.5: An example illustrating the reason for straight median tracks. The gray line with noise indicates a portion of a raw median track extracted from an ensemble; the yellow line shows the smoothed path. The red points represent the sample points of this ensemble over a period of time, indicating no samples at the right side of the smoothed median track.

Least Squares-based method described in Chapter 4, initiating the RBF interpolation system based on this subset. Doing so enables the assigning of a value of storm intensity to an arbitrary location covered by the spatial spread of the ensemble at time $j$. To interpolate storm intensities over the complete time period of the forecasting, RBF interpolation systems for all points in time are constructed. Other storm characteristics such as storm size are subsequently incorporated using the same strategy.

## 5.2 Applications to Additional NHC Ensembles

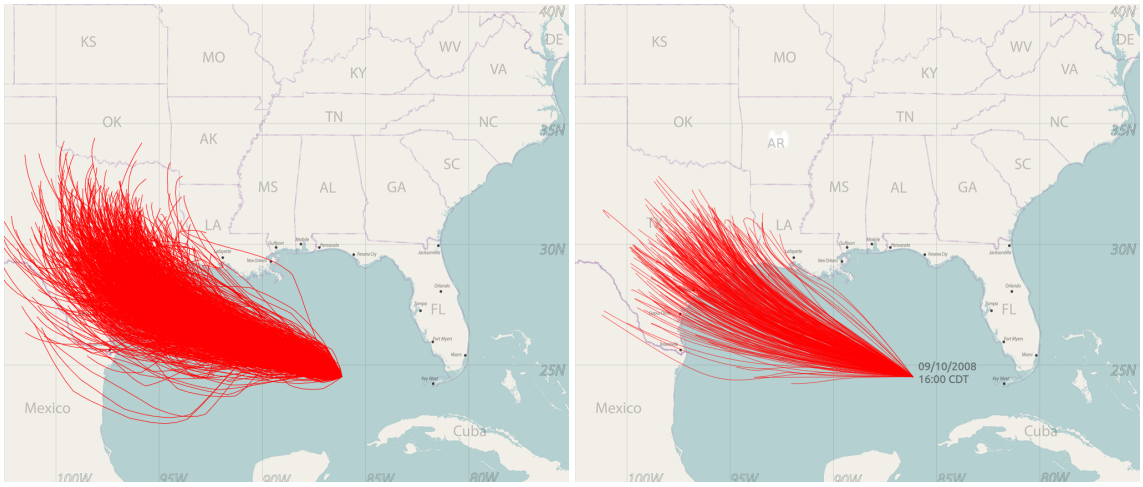
This section reports the experimental results of using the algorithmic framework developed here to extract small, representative subsets of the predicted tracks of the following hurricanes from their respective NHC forecast ensembles: Hurricane Katrina on August 27, 2005, Hurricane Rita on September 22, 2005, Hurricane Ike on September 10, 2008, Hurricane Ida on December 8, 2009, and Hurricane Issac on August 26, 2012. Each of these ensembles comprises 1,000 member tracks, with hourly time-sampled predicted locations (latitude and longitude), storm sizes, and storm intensities over the subsequent 120 hours from the beginning of the predictions. For demonstration purposes, the work reported is limited to 72-hour predictions from the

track ensembles; however, this technique can be applied to complete data sets without any modification.

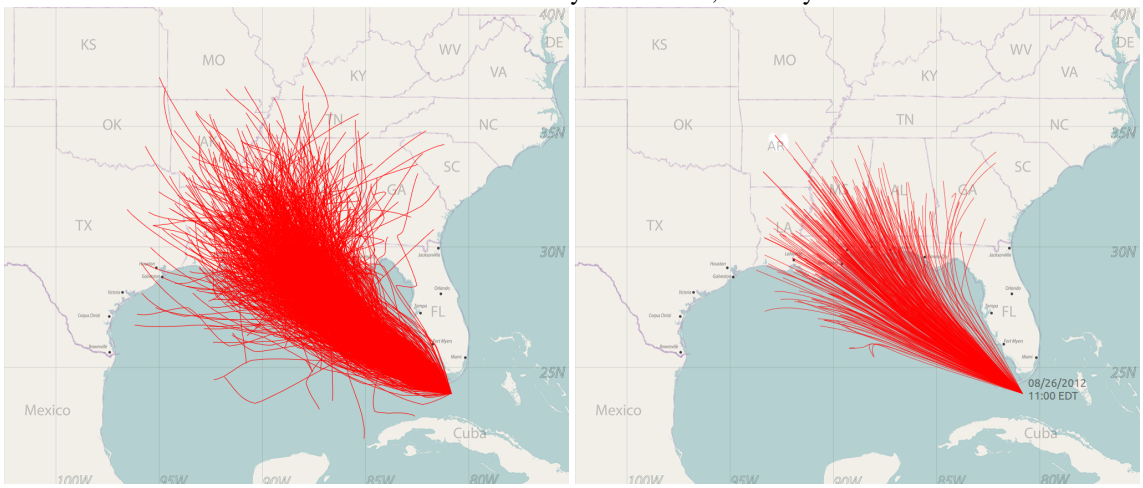
To verify that the representative tracks extracted using the approach proposed in this research accurately preserve the spatial distribution included in the original predictions, the original ensembles of these hurricanes and their corresponding subsets of representative tracks are compared as demonstrated in Figure 5.6, the first column of which shows the plots of the ensembles, while the second shows their corresponding subsets of median tracks, the sizes of which are maximized to avoid the inaccurate straight paths described in the previous section. This comparison first reveals that such subsets preserve the most significant trends of these hurricanes because the high densities of tracks covered by the subsets seen in the center regions are nearly identical to those covered by the original ensembles. In addition, the individual subsets exhibit approximately identical spatial spreads along the perpendicular directions of the forward directions of the hurricanes to those included in the original ensembles.

However, the spatial coverages of these representative subsets are shorter than those of the original ensembles along the forward directions, meaning that they potentially underestimate the uncertainty of the predicted speeds of the hurricane. This issue is more clearly seen when details of the coverages and their distributions are compared at specific points in time. For instance, Figure 5.7 exemplifies such details for Hurricane Ida 2005 at 60 and 72 hours from the beginning of the prediction: the red points represent the predicted locations extracted from the original ensemble, the simplicial depths of which are color-encoded using luminance of red, meaning that the color of a deeper location tends to be brighter; the gray polylines represent representative tracks extracted from the ensemble using the proposed approach; and the blue points represent locations along the representative tracks at these specific points in time. As these figures show the blue points preserve the center of the distribution represented by the red as they cover the area with the brightest red points, and they indicate the most significant spatial uncertainty of the width of the distribution of the red points because the blue outliers closely match those of the red along the direction perpendicular to the forward direction of the hurricane. However, they cannot maintain the distribution along the elongation of the forward direction of the hurricane, a result of the variation of the predicted speed of the hurricanes included in the forecasts.

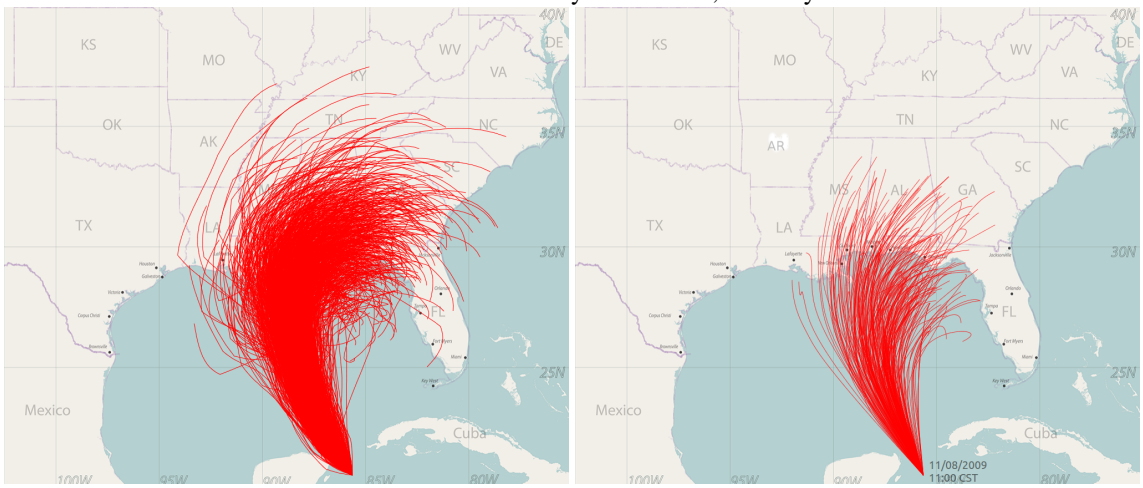
More specifically, such variations can potentially be visualized by drawing multiple points along the hurricane path to indicate a probabilistic spatial range of the predicted location at a particular point in time as shown in Figure 5.8(a), where the black line represents a segment of a representative track, the solid blue circles represent the most certain predicted location at a specific point in time, and the dashed blue circles represent other possibilities of this predicted location, reflecting the uncertainty of the predicted hurricane



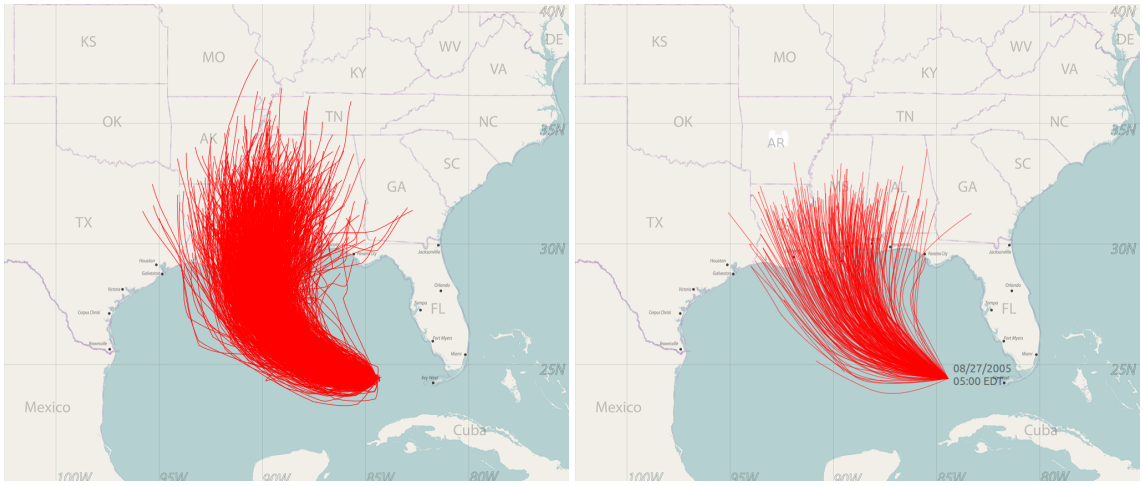

Ike: 4PM CDT Wednesday 09/10/2008, advisory 39



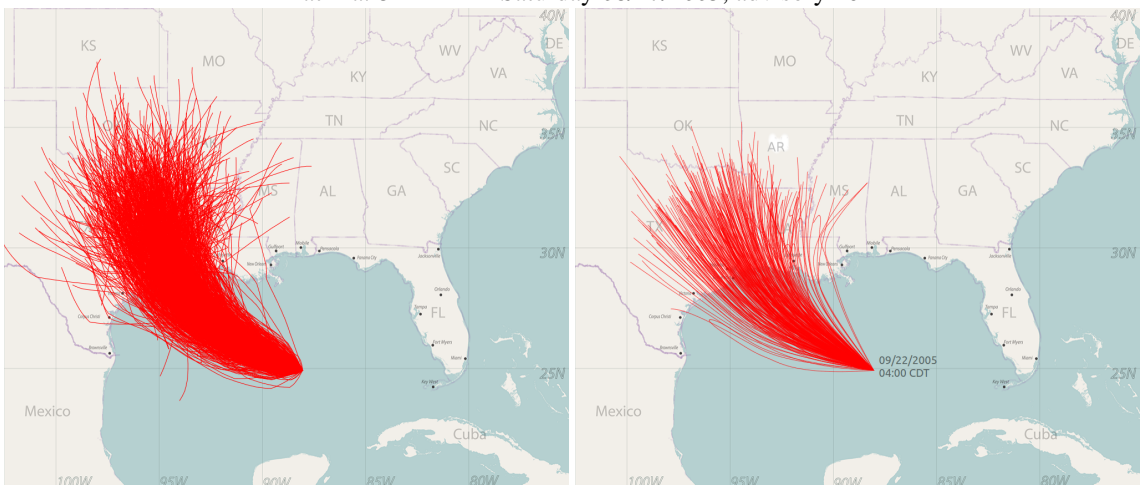

Isaac: 11 AM EDT Sunday 08/26/2012, advisory 22




Ida: 3AM CST Sunday 11/08/2009, advisory 17

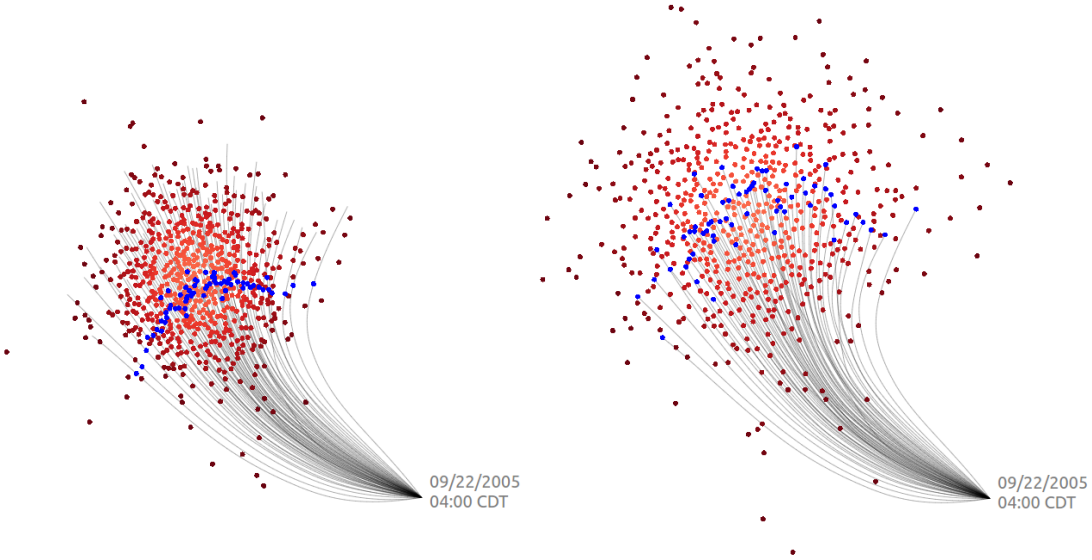



Katrina: 5AM EDT Saturday 08/27/2005, advisory 16



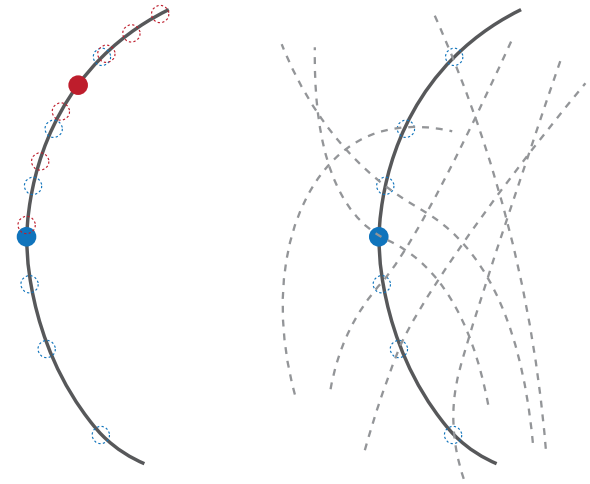

Rita: 4AM CDT Thursday 09/22/2005, advisory 17

Figure 5.6: A comparison between the original forecast ensembles of 5 hurricanes (the first column) and their corresponding subset of representative tracks extracted using the approach proposed in this research (the second column).




(a) (b)

Figure 5.7: Detailed illustrations of representative tracks of Hurricane Ida 2005 at 60 (a) and 72 (b) hours from the beginning of the prediction.



(a) (b)

Figure 5.8: Two possible approaches to visualize the uncertainty of predicted hurricane speeds which inevitably lead to visual clutter.

speed. However, these circles will inevitably overlap with those portraying possible predicted locations of adjacent hours as represented by the red circles in this figure. Another approach for including these variations is to compute more representative forecast tracks passing through those possible locations, a method that can significantly increase the number of tracks, resulting in the visual clutter shown in Figure 5.8(b) where these additional tracks are represented by dashed polylines. Therefore, this research posits that the uncertainty of hurricane speed is difficult to accurately preserve in such static visualizations using polylines to represent hurricane forecast tracks and using the layout of these polylines to implicitly portray the predicted spatial uncertainty included in the original forecast ensembles. This limitation, however, can be effectively addressed using the other point-based visualizations developed in Chapter 4.

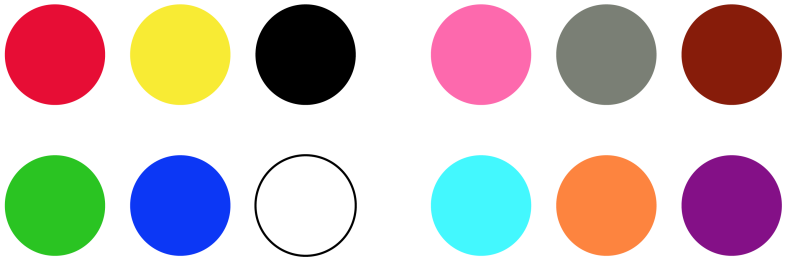
## 5.3 Visualization Design

To determine the size of these track subsets for the visualization proposed here, two requirements are considered. First, a subset should be able to cover as much as possible the spatial spread of the full ensemble to provide viewers enough information to interpret the forecast. Second, since an NHC track ensemble usually retains a higher density near the regions of its spatial center, adding too many members to the extracted subset can lead to visual confusion such that additional manipulations, e.g. varying colors or changing shapes, used for visually encoding ancillary storm characteristics are difficult to be perceived by

viewers. To satisfy these conditions, the work reported here produces subsets comprised of 15 tracks that convey approximately 87.5% spatial spread of the full ensembles. The size of this subset can be adjusted for applying the technique to different types of simulation ensembles.

Subsequently, the subsets obtained from the proposed sampling framework are used as basis for building direct ensemble visualizations. Such visualizations portray extracted median paths using polylines such that the predicted spatial information and its uncertainty included in the original ensemble are implicitly depicted by the spatial layout of these polylines. Typically, the more certain an area is predicted to experience the hurricane, a higher concentration of median paths is exhibited in this area, and vice versa.

The predicted storm intensities are encoded in this visualization by rendering a segment of a median path using a color determined by the predicted intensity attached to the starting location of this segment. The NHC ensembles describe storm intensity as the maximum predicted wind speeds measured in nautical miles per hour. To make it more intuitively understandable for the general public, these wind speeds are converted to a widely adopted storm taxonomy system, known as the Saffir-Simpson Hurricane Wind Scale [NOA17d], categorizing hurricane intensities as tropical depression and/or storm, and Category 1 to 5 hurricanes. By doing so, storm intensities are mapped to seven levels that can be most appropriately portrayed using a diverging color scheme. Additionally, because the intensity of a hurricane usually changes gradually over time and hourly sampled intervals of a hurricane track can be short in the display, the color scheme used in this visualization should be able to make adjacent hurricane categories more visually distinguishable. To satisfy these requirements, the diverging color scheme recommended by Ware [War04] shown in Figure 5.9(a), which has the advantage of using colors that are easily remembered and recognized, is consulted here. In this research, blue, cyan, white and black are intentionally avoided as they are too close to the colors used in the graphical elements and texts composited in the background map, making them difficult to be distinguished by the viewers. Yellow is also not selected as it is so bright that it can be hardly interpreted when displaying it on the white-based background. Therefore, the remaining seven colors, i.e. red, green, pink, gray, orange, brown, and purple, are employed in this visualization. They are slightly modified to prevent any individual color from being over-emphasized and make them more clearly perceived by viewers. The RGB values of these modified colors are: (90, 174, 97) for green, (102, 102, 102) for gray, (153, 112, 171) for purple, (191, 129, 45) for brown, (224, 130, 20) for orange, (255, 127, 127) for pink, (255, 80, 80) for red. These colors are sorted in an ascending order based on their values of the red channel to indicate an increase in the intensity, meaning that the green represents the safest situation while the red depicts the most severe level of risk. The resulted color scheme is illustrated in Figure 5.9(b), where the tropical depression, tropical storm, and



(a) A suggestion by Colin Ware[War04]

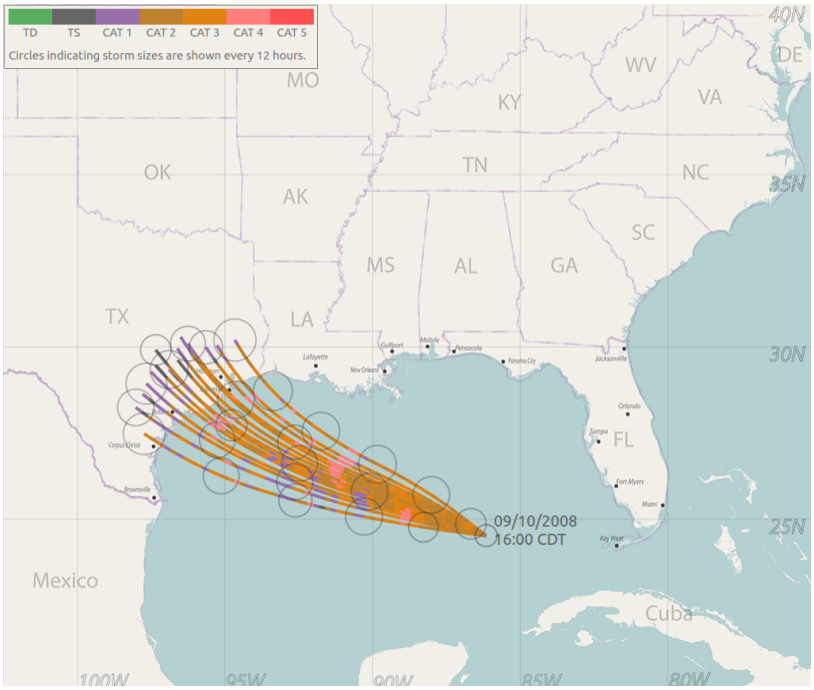

(b) The color scheme developed in this research.

Figure 5.9: Diverging color schemes.

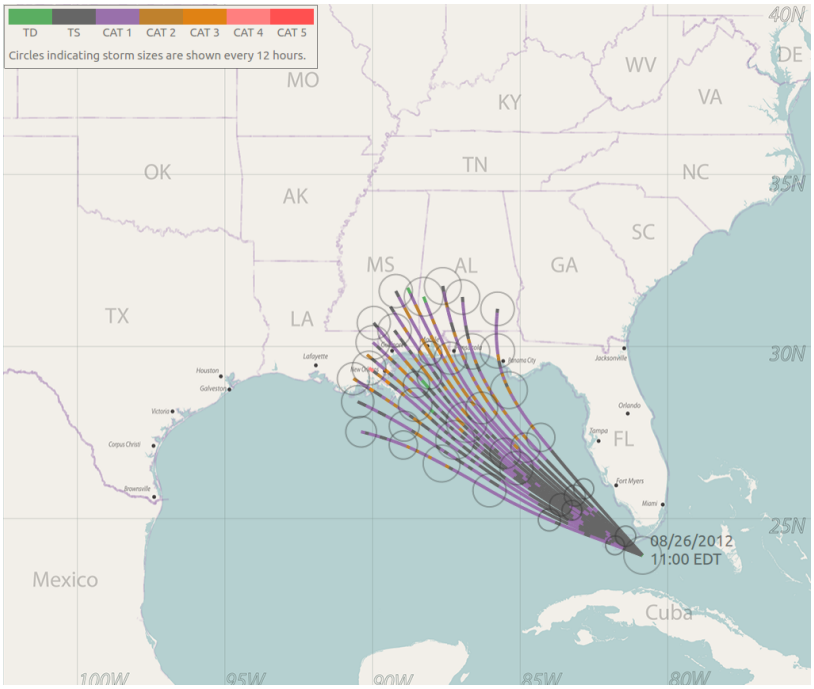
Category 1 to 5 hurricanes are mapped to the green, gray, purple, brown, orange, pink, and red, respectively.

Finally, the predicted storm sizes are included in this visualization by adding circles centered at locations along the tracks. The radii of these circles are determined by the predicted radii of the 64-kt wind at these locations, thereby the areas encompassed by these circles imply possible regions predicted to experience the hurricane. Since drawing storm size circles at every hour of a prediction can easily lead to visual clutter, they are shown for every 12 hours. In addition, since hurricane forecasts have less uncertainty during early lead times, resulting in tracks that are highly concentrated in a smaller region, the circles of tracks during these lead times can also be crowded. To address this issue, for individual time points, circles belonging to the most left and right tracks are kept, while the others that are too close to another are not shown. By doing so, the area potentially affected by hurricane at a certain level of confidence is depicted using a few easily visible and interpreted circles. The resulted visualizations for the five NHC forecast track ensembles are illustrated in Figure 5.10.

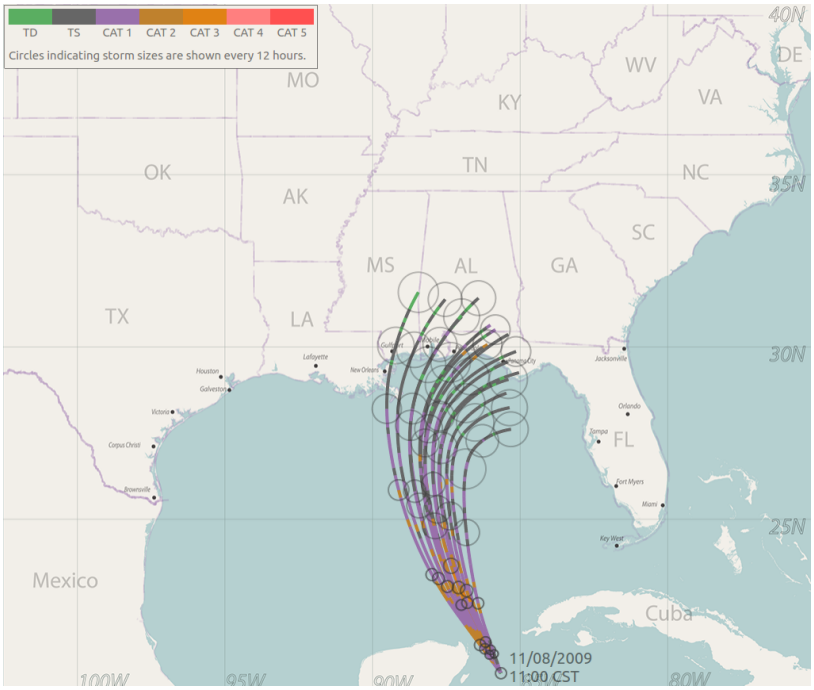
Another important benefit of this visualization is that the crucial time information included in a forecast can be clearly depicted. For instance, as Figure 5.10(d) shows, the hurricane forecast as a Category 3 was predicted to develop to a Category 4 approximately 36 hours from the beginning of the initial prediction. Then it would decrease to a Category 2 or 3 before landing on the coastline at 60 hours and diminishing to a Category 1 hurricane or tropical storm over the next 12 hours. According to a communication with the FEMA representative in the NHC, such time information is needed to support emergency decision-making when a hurricane is approaching and all visualizations published by the NHC, including the widely used cone



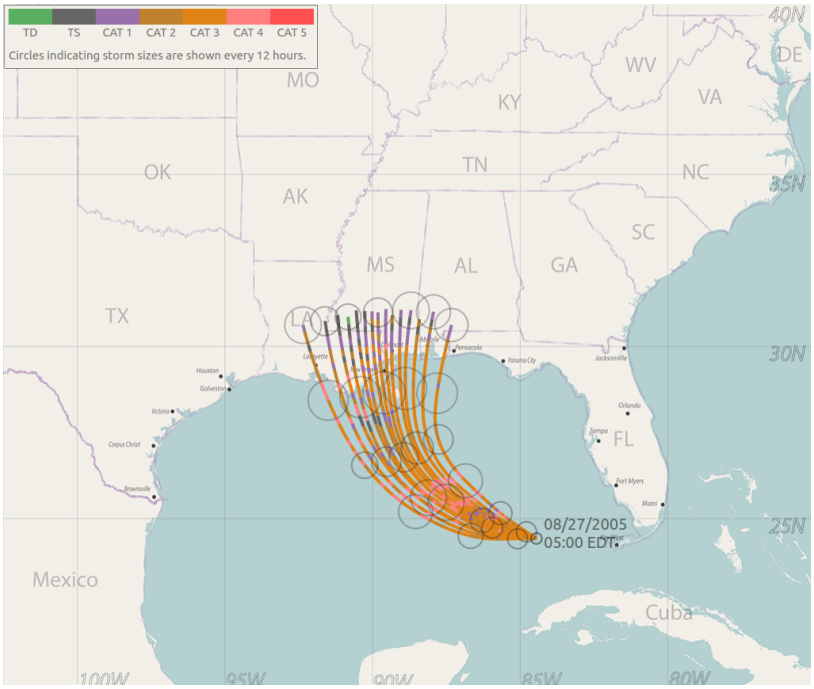

(a) Ike: 4PM CDT Wednesday 09/10/2008, advisory 39



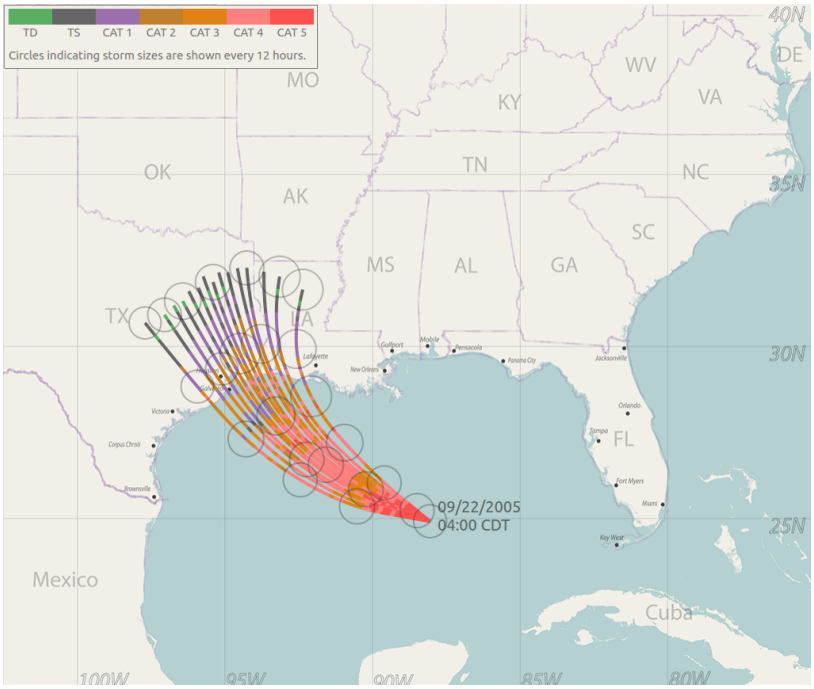

(b) Isaac: 11 AM EDT Sunday 08/26/2012, advisory 22




(c) Ida: 3AM CST Sunday 11/08/2009, advisory 17




(d) Katrina: 5AM EDT Saturday 08/27/2005, advisory 16




(e) Rita: 4AM CDT Thursday 09/22/2005, advisory 17

Figure 5.10: Developed visualizations of the NHC forecast track ensembles of five hurricanes.

of uncertainty, cannot accurately portray this information.

## 5.4 Experimental Evaluation

To evaluate the effectiveness of the visualization of representative hurricane forecast tracks developed here in enhancing viewers' estimates of hurricane risk, a subsequent cognitive study is being conducted in collaboration with researchers from the University of Utah. This study adopts the experiment conducted for the time- and location-specific visualizations (Section 4.4), focusing on whether the judgments of damage made by the viewers and their confidence levels in them varied as a function of visualization type, distance from the center of the forecast, and forecasting time.

More specifically, this study compares four types of displays generated using the representative sampling framework developed in this research. They are categorized as tracks-only display and annotated visualization, where the former shows different numbers of hurricane forecast tracks using red polylines, while the latter is the visualization developed consisting of 15 tracks including annotations of the predicted storm sizes and intensities. Figure 5.11(a), (b), and (c) exemplify the tracks-only displays of 7, 15, and 63 tracks,

respectively, and (d) exemplifies the annotated visualization. These displays are generated for the five NHC ensembles introduced in the previous section. To reflect the variation of distance from the center of the forecast, a location of a simulated offshore oil rig platform is superimposed on individual displays, as exemplified by the blue dot in Figure 5.11, at one of eight locations defined relative to the spatial center of the prediction (0.0) and the NHC official forecast error (1.0). These locations are determined as $\pm 0.2, 0.8, 1.10, 1.78$ with respect to the forecast error. For instance, if the official NHC forecast error is $R$ km, the simulated oil rig platform is set at $\mu R$ km away from the center track of the forecast along the perpendicular direction of the hurricane forward direction, where $\mu$ is one of these fractions. To incorporate the information concerning the forecasting time, these locations are generated at both 24 and 48 hours into the prediction, where the forecast errors adopted here were 186.07 km at 24 hours and 347.28 km at 48 hours.

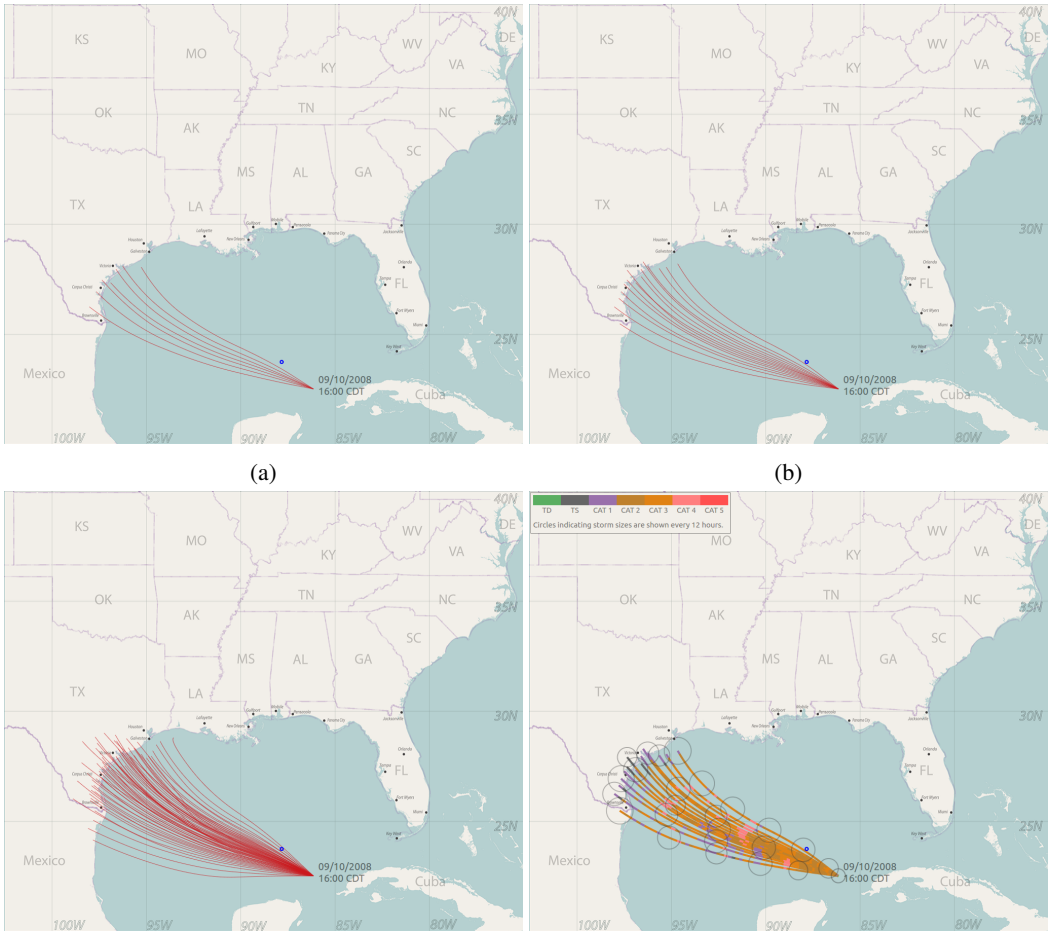
After being given a description of the oil rig platform, each participant is shown one of the four types of displays and completes 80 trials in which the locations of the platform (8 conditions), the forecasting time points (2 conditions), and the hurricanes (5 conditions) are randomly presented. In each trial, the participants are instructed to estimate the level of damage that would incur at the simulated oil rig based on the likelihood of the storm affecting it and its strength in the affected region. They are also instructed to evaluate the level of confidence of their estimations. A screen-shot of the experiment can be seen in Figure 5.12, participants are provided with a Likert scale ranging from 1 (no damage or not at all confident) to 7 (severe damage or very confident) for entering their answers. Finally, they are asked to describe their decision-making process in a text box and answer a series of six true/false statements to assess their understanding and misconceptions of the displays as shown in Table 5.1.

Table 5.1: A list of true/false statements provided in the cognitive study.

| Statements |
|---|
| 1. Do you have experience with hurricane forecasts? |
| 2. Have you lived or do you live in an area that experiences hurricane threats? |
| 3. The display shows the hurricane getting large over time. |
| 4. The display indicates that the forecasters are less certain about the path of the hurricane as time passes. |
| 5. Areas on the map not covered by the visualization will not be hit by the hurricane. |
| 6. Did you recognize the areas in the map? |

### 5.4.1 Preliminary Experimental Results

As this study is still ongoing, only preliminary results analyzed from current data collected from 56 participants are reported here.




(a) (b)

(c) (d)

Figure 5.11: Examples of visualizations compared in the cognitive study: (a), (b), and (c) show 7, 15, and 63 median tracks without annotations including additional information concerning the predicted storm sizes and intensities, respectively, while (d) shows 15 median tracks with annotations of these storm characteristics. The blue dots indicate the position of a simulated oil rig platform.

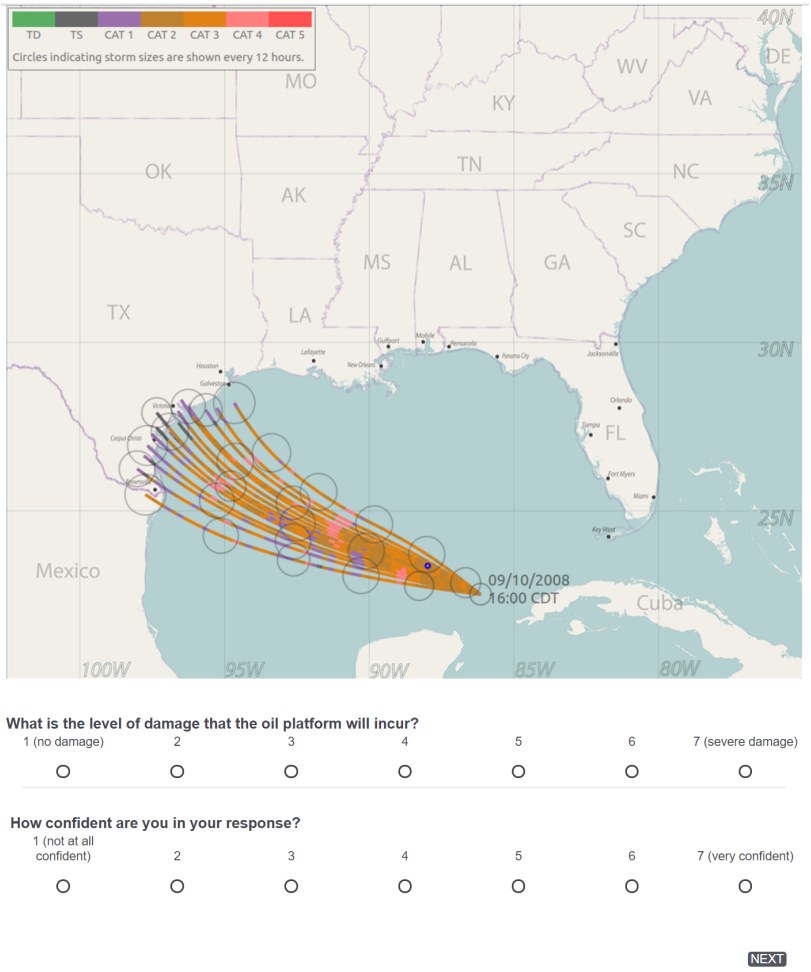



Figure 5.12: A screen-shot of the experiment testing the effectiveness of visualizations of representative hurricane forecast tracks.

#### 5.4.1.1 Data Analysis Approach

Multilevel models were fitted to the data using *R* with the *lme4* package [BMBW15] and using restricted maximum likelihood estimation procedures [RABCdT11, RB02]. A multilevel model is appropriate for modeling these data and, thus, for testing the hypotheses in this study for two primary reasons: 1) it allows for the inclusion of interactions between continuous variables (in this study, distance) and categorical predictors (in this study, the type of visualization and time point); 2) it uses robust estimation procedures appropriate for partitioning variance and error structures in mixed and nested designs (repeated measures nested within individuals in this case).

More specifically, the mixed two-level regression model was used here to test whether the judgments of damage made by the viewers and their confidence level in them varied as a function of visualization type (level 1) and distance from the centers of the forecast and forecasting time points (level 2). Visualization type was coded such that the participant group shown the 7-tracks-only display was the referent in the comparison among the number of tracks, while the group shown 15-tracks-only was the referent in the comparison between displays with and without annotations. Mathematically, the first level of this regression model for modeling estimated damage is

$$\text{Damage}_{ij} = \beta_{0j} + \beta_{1j} \times \text{Distance}_{ij} + \beta_{2j} \times \text{Time}_{ij} + \beta_{3j} \times \text{Distance}_{ij} \times \text{Time}_{ij} + r_{ij}$$

while the second level is

$$\begin{aligned}
\beta_{0j} &= \gamma_{00} + \gamma_{01} \times \text{Visualization}_j + u_{0j}, \\
\beta_{1j} &= \gamma_{10} + \gamma_{11} \times \text{Visualization}_j + u_{1j}, \\
\beta_{2j} &= \gamma_{20} + \gamma_{21} \times \text{Visualization}_j + u_{2j}, \\
\beta_{3j} &= \gamma_{30} + \gamma_{31} \times \text{Visualization}_j + u_{3j},
\end{aligned}$$

where $i$ represents trials, $j$ represents individuals, the $\beta$ and $\gamma$ terms are the regression coefficients, and the

$r_{ij}$ and the $u$ terms are error terms. These equations can be rewritten in a mixed format:

$$\begin{aligned}\text{Damage}_{ij} = {} & \gamma_{00} + \gamma_{01} \times \text{Visualization}_j + \gamma_{10} \times \text{Distance}_{ij} + \gamma_{20} \times \text{Time}_{ij} + \\ & \gamma_{11} \times \text{Visualization}_j \times \text{Distance}_{ij} + \gamma_{21} \times \text{Visualization}_j \times \text{Time}_{ij} + \gamma_{30} \times \text{Distance}_{ij} \times \text{Time}_{ij} + \\ & \gamma_{31} \times \text{Visualization}_j \times \text{Distance}_{ij} \times \text{Time}_{ij} + \\ & u_{0j} + u_{1j} \times \text{Distance}_{ij} + u_{2j} \times \text{Time}_{ij} + u_{3j} \times \text{Distance}_{ij} \times \text{Time}_{ij} + r_{ij}. \end{aligned} \tag{5.2}$$

Similarly, the model for examining participants' confidence levels in their damage estimations is:

$$\begin{aligned}\text{Confidence}_{ij} = {} & \gamma_{00} + \gamma_{01} \times \text{Visualization}_j + \gamma_{10} \times \text{Distance}_{ij} + \gamma_{20} \times \text{Time}_{ij} + \\ & \gamma_{11} \times \text{Visualization}_j \times \text{Distance}_{ij} + \gamma_{21} \times \text{Visualization}_j \times \text{Time}_{ij} + \gamma_{30} \times \text{Distance}_{ij} \times \text{Time}_{ij} + \\ & \gamma_{31} \times \text{Visualization}_j \times \text{Distance}_{ij} \times \text{Time}_{ij} + \\ & u_{0j} + u_{1j} \times \text{Distance}_{ij} + u_{2j} \times \text{Time}_{ij} + u_{3j} \times \text{Distance}_{ij} \times \text{Time}_{ij} + r_{ij}. \end{aligned} \tag{5.3}$$

#### 5.4.1.2 Results

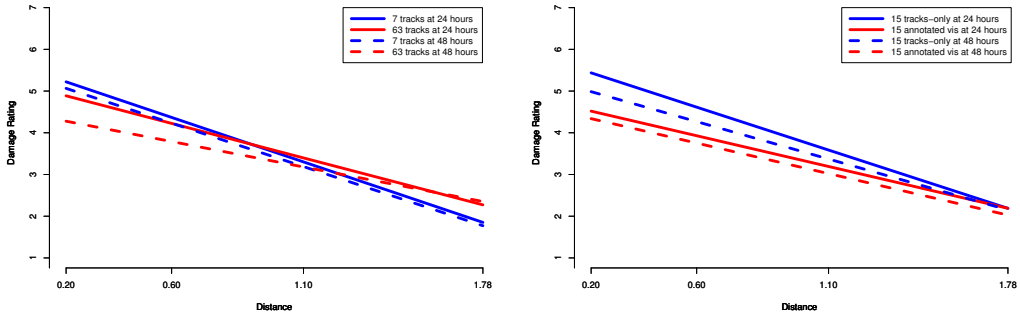
Table 5.2 shows the estimated regression coefficients which tests whether the judgments of damage varied as a function of the number of representative forecast tracks, distance from the center of the forecast, the forecasting time points, and their interactions, where *VisNumD*1 and *VisNumD*2 represent 15- and 63-tracks-only displays. First, as the $p-value$ of $Distance \times VisNumD2$ (0.003) is less than 0.05, meaning that this interaction is statistically significant at a 5% significance level. This analysis reveals that the association between the distance from the oil rig platform to the spatial center of the prediction (distance-to-center) and the participants' judgments of damage significantly changed when the number of tracks increased from 7 to 63, when averaged across both time points. The significant interaction between $Time \times VisNumD2$ suggests that the effect of time point is different for the 7-tracks-only display compared to the 63-tracks-only display. There appears to be little to no effect of time point for the 7-tracks-only display and a larger effect for the 63-tracks-only display. These results can be seen in Figure 5.13, where the blue and red solid lines represent the results of placing the oil rig platform at 24 hours in the prediction for the 7- and 63-tracks-only displays, while the blue and red dashed lines represent the results for placing this platform at 48 hours for the 7- and 63-tracks-only displays. As shown in Figure 5.13(a), the red dashed line is flatter than the others, meaning that when shown the 63-tracks-display at 48 hours, the participants were less likely to associate the potential

Table 5.2: Estimated MLM regression which tests whether the judgments of damage varied as a function of the number of the representative forecast tracks, distance from the center of the forecast, and the forecasting time points.

| *Fixed Effect* | *Estimate* | *St. Error* | *t-stat* | *p-value* | |
|---|---|---|---|---|---|
| Intercept ($\gamma_{00}$) | 5.647 | 0.253 | 22.310 | 0.000 | *** |
| Distance ($\gamma_{10}$) | -2.132 | 0.106 | -20.117 | 0.000 | *** |
| Time ($\gamma_{20}$) | -0.166 | 0.163 | -1.015 | 0.310 | |
| VisNumD1 ($\gamma_{01}$) | 0.199 | 0.372 | 0.536 | 0.591 | |
| VisNumD2 | -0.430 | 0.392 | -1.097 | 0.272 | |
| Distance×Time ($\gamma_{30}$) | 0.048 | 0.149 | 0.325 | 0.744 | |
| Distance×VisNumD1 ($\gamma_{11}$) | 0.076 | 0.156 | 0.491 | 0.622 | |
| Distance×VisNumD2 | 0.478 | 0.164 | 2.910 | 0.003 | ** |
| Time×VisNumD1 ($\gamma_{21}$) | -0.336 | 0.241 | -1.396 | 0.162 | |
| Time×VisNumD2 | -0.529 | 0.253 | -2.087 | 0.036 | ** |
| Distance×Time×VisNumD1 ($\gamma_{31}$) | 0.212 | 0.220 | 0.964 | 0.335 | |
| Distance×Time×VisNumD2 | 0.387 | 0.232 | 1.667 | 0.095 | * |

*Note*: *, **, *** indicates that a null hypothesis of zero is rejected at the 10%, 5%, and 1% level, respectively.

damage to the distance from the center of the prediction and, thus, comprehend the spatial spread of the tracks as an uncertainty distribution rather than incorrectly interpreting the center as the region predicted to experience the most intense storm wind. However, the results comparing 15 to 7 tracks did not show any statistical significance.




(a) (b)

Figure 5.13: Plots of estimated MLM regression which tests whether the judgments of damage varied as a function of the number of the representative forecast tracks, distance from the center of the forecast, and the forecasting time points: (a) shows the result comparing the 63- to 7-tracks-only displays and (b) shows the result comparing the annotated visualization to the 15-tracks-only display.

Table 5.3 shows the analysis comparing the displays of 15 tracks with and without annotations. As this analysis reveals, using annotations can significantly influence the relationship between distance-to-center and damage rating as the $p-value$ of $Distance \times Annotation$ is less than 0.001. Typically, including

annotations in a display results in lower damage ratings at the spatial center of the prediction; and as they change gradually as the distance-to-center increases, this annotated visualization is comprehended more like an uncertainty distribution. Even though using this annotated visualization was not directly statistically compared to using the 63-tracks-only display in the regression model, a comparison between their trend lines, as shown in Figure 5.14, indicates that the results of using the annotated visualization at the 24- and 48-hour are approximately identical to that of using the 63-tracks-only at 48 hours.

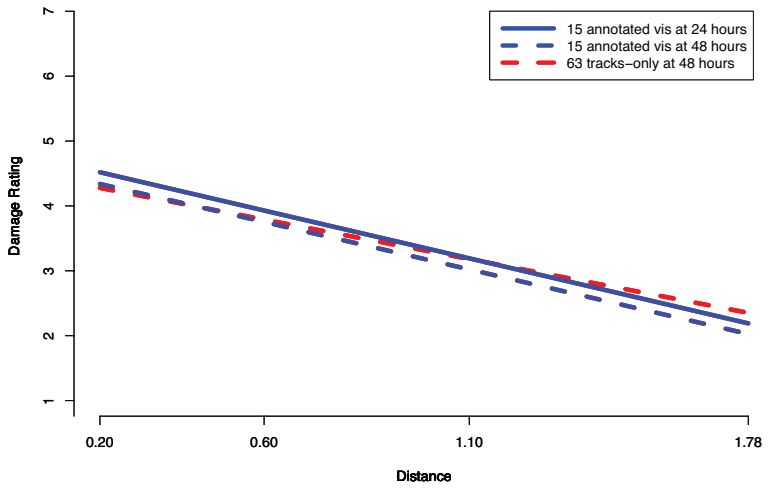



Figure 5.14: A comparison between using the 63-tracks-only display and using the 15 annotated visualizations.

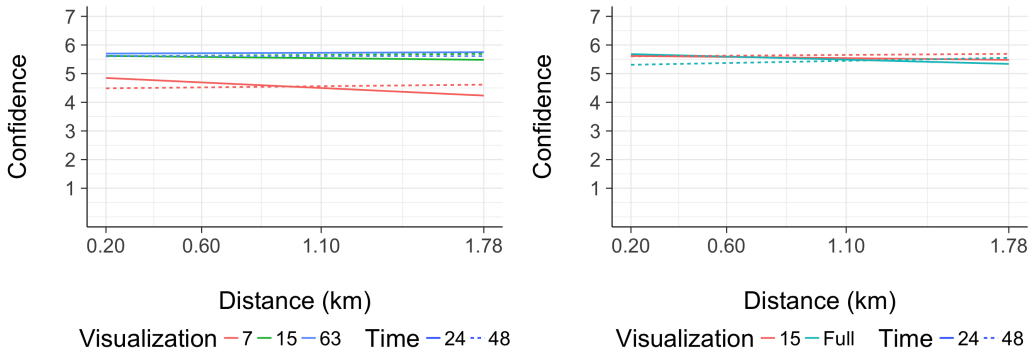
Figure 5.15 plots the regression lines between the relationship of distance-to-center and the participants' confidence level in their judgments of damages when using different visualizations at different

Table 5.3: Estimated MLM regression which tests whether the judgments of damage varied as a function of the visualization with or without annotations, distance from the center of the forecast, and the forecasting time points.

| *Fixed Effect* | *Estimate* | *St. Error* | *t-stat* | *p-value* | |
|---|---|---|---|---|---|
| Intercept ($\gamma_{00}$) | 5.847 | 0.304 | 19.198 | 0.000 | *** |
| Distance ($\gamma_{10}$) | -2.056 | 0.111 | -18.398 | 0.000 | *** |
| Time ($\gamma_{20}$) | -0.503 | 0.172 | -2.912 | 0.003 | *** |
| Annotation ($\gamma_{01}$) | -1.034 | 0.385 | -2.684 | 0.007 | *** |
| Distance×Time ($\gamma_{30}$) | 0.261 | 0.158 | 1.655 | 0.097 | * |
| Distance×Annotation ($\gamma_{11}$) | 0.582 | 0.141 | 4.118 | 0.000 | *** |
| Time×Annotation ($\gamma_{21}$) | 0.321 | 0.218 | 1.469 | 0.141 | |
| Distance×Time×Annotation ($\gamma_{31}$) | -0.251 | 0.199 | -1.256 | 0.208 | |

*Note*: *, **, *** indicates that a null hypothesis of zero is rejected at the 10%, 5%, and 1% level, respectively.

forecasting time points. Figure 5.15(a) shows that participants were less confident about their estimations when shown the 7-tracks-only display than the 15- and 63- tracks-only displays. There is no significance between using 15 and 63 tracks. Figure 5.15(b) shows that when participants were shown the annotated visualization, they felt as confident in their estimations as they did when shown the 15-tracks-only visualizations.




(a) (b)

Figure 5.15: Plots of the estimated MLM regression which tests whether the participants' confidence levels in their judgments of damage varied as a function of the type of visualization, distance from the center of the forecast, and the forecasting time points.: (a) shows the result comparing the 63- and 15- to 7-tracks only displays and (b) shows the result comparing the annotated visualization to 15-tracks-only display.

### 5.4.2 Experiment Summary

Overall, this experimental reveals that when shown 63 tracks at late leading time (48 hours), the participants were less likely to incorrectly associate the damage of the storm to the spatial center of the forecast than when shown 7 and 15 tracks. However, when shown 15 tracks including the annotations of the storm sizes and intensities developed in this research, participants achieved a similar accuracy as when shown 63 tracks and additionally, this accuracy was preserved at both early (24 hours) and late (48 hours) lead times. These results suggest that the annotated visualization is comprehended more precisely as an uncertainty distribution and thus, is an asset in enhancing viewers' hurricane emergency decision-making. As this experiment is still ongoing and these results are analyzed from preliminary data collected from a small number of participants (56), a more complete analysis of the data collected from a larger number of participants (200) will be conducted in the future to more fully validate the effectiveness of the annotated visualization.

# Chapter 6

# Conclusion

## 6.1 Summary

When a hurricane is approaching, the questions that most concern the local governmental officials, news media, and the general public are when and where will it make landfall and its intensity. These questions cannot be accurately answered by using the NHC track forecast cone, which is the most widely used graphic product for displaying forecasts, because it does not appropriately represent such storm characteristics, nor does it represent their uncertainties. Therefore, viewers have difficulty making appropriate emergency decisions, either resulting in unnecessary expenditures or more damage and deaths. The work reported here addresses these issues by developing visualizations that improve the accuracy in depicting hurricane forecasts and their uncertainties.

This research began by more effectively portraying the spatial-temporal information included in hurricane predictions than current visualizations of hurricane forecasting can. To address this challenge, a summary display rendering three levels of positional storm strike risk at a specific point in time was created based on the development of a RBF-based interpolation technique. While this visualization effectively incorporates the important time information, it potentially leads to misconceptions of the prediction and its uncertainty because of the intrinsic limitation of using summary displays. To address this limitation, direct ensemble visualizations, which explicitly render members of hurricane forecast ensembles and, thus, implicitly convey the uncertainty using the spatial spread of these members, were developed. To avoid the visual clutter observed in the visualization due to the large size of the ensembles, this research developed a representative sampling framework for processing 2D points by extracting small subsets from the origi-

nal ensembles. To depict the uncertainties of storm size and intensity, two animated visualizations of the subsets obtained were created, with one representing the areas affected by a hurricane using solid circles at the hurricane-predicted locations and the other one representing the intensity by drawing NHC glyphs at these locations. Even though the effectiveness of these visualizations has been supported by a subsequent cognitive study, depicting the complete information of a hurricane prediction over a hurricane forecast time period requires producing and displaying a large number of animations, thus making it difficult for viewers to fully comprehend the prediction. To address this difficulty, this research generalized the sampling framework by processing ensembles of hurricane-predicted tracks, enabling a direct ensemble visualization displaying a representative subset of these tracks. This visualization not only represents the predicted storm characteristics including locations, storm sizes, and storm intensities but also reveals the significant time information included in hurricane forecasts. An ongoing preliminary cognitive study examining the effectiveness of this visualization has suggested that it minimizes the confusion concerning the spatial attributes of a prediction with the underlying uncertainty, and thus, it has the potential to enhance emergency decision-making.

## 6.2 Extensions and Future Research Directions

Even though the goal of the techniques developed in this research was to produce novel uncertainty visualizations for hurricane forecasting, they can potentially be extended to benefit the field of ensemble visualization more generally. For example, the sampling frameworks used for extracting small, representative subsets from hurricane forecast ensembles can potentially be generalized to address the challenge of large ensemble sizes. As the size and complexity of these ensembles increase continuously, it is difficult to portray the large amount of information included in them effectively so that they can be easily and accurately interpreted by viewers. One potential research direction is to explore the ability of these sampling frameworks developed here to improve scatter plots, a method widely adopted by the visualization community for depicting multivariate ensembles and the included uncertainty, by addressing the over-plotting resulting from large ensembles. More specifically, as a scatter plot visualizes a bivariate pair of variables included in a dataset by drawing glyphs at Cartesian coordinates, which are inevitably positioned on top of one another as the size of the dataset grows, making it difficult for the viewers to see individual glyphs, subsequently impacting the analysis of the data. The proposed technique can significantly reduce the number of data samples needing to be displayed by selecting a small subset from the original large dataset, the spatial distribution of which is accurately preserved. More importantly, as the crucial uncertainties are usually incorporated into the scatter

plot by changing visual attributes of the glyphs, e.g. color, size, and shape, as exemplified in Figure 6.1, this technique provides more flexibility for such incorporations as members of the subset are guaranteed to achieve a spatial minimum between one and another, making it easier for viewers to perceive these variations.

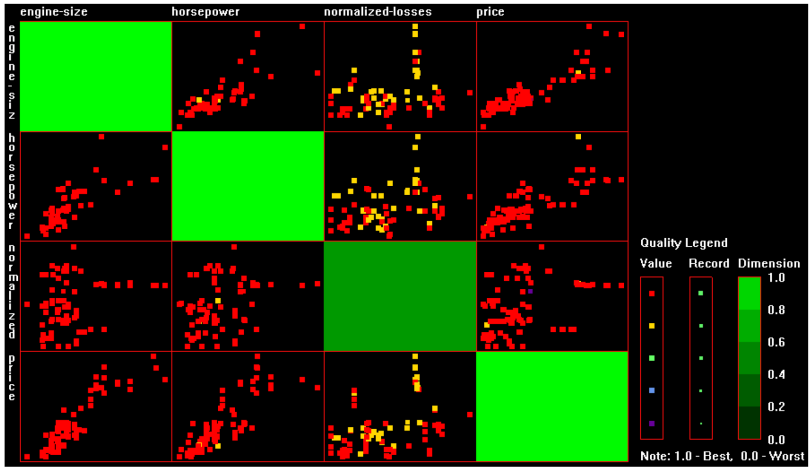



Figure 6.1: A scatter plot matrix of a dataset of automobile specifications, where qualities of data value, record, and dimension reflecting various uncertainties of the dataset are visually encoded using hue, size, and saturation, *Source*: [XHWR06].

In addition to improving the standard scatter plots of ensembles, the proposed sampling approach can potentially enhance density plots, an alternative technique developed for visualizing large ensembles without the over-plotting problem. Typically, a density plot depicts an ensemble by estimating and portraying its probabilistic density function (PDF), the regions with high densities being emphasized in the visualization, as exemplified in Figure 6.2. An important limitation of this technique is that the size of the dataset may become too large for the efficient estimation of its PDF. The current solution for addressing this limitation computes the density field using a small collection of randomly generated data samples, the distribution of which is approximately identical to that of the original. However, this technique is still limited because current approaches for sampling a distribution are not applicable for ensembles with unknown PDFs [FKLT10], which are frequently seen in the real world. The approach developed in this research, which does not require an *a priori* assumption of the PDF of the ensemble, has the potential to address this problem. In addition, this approach selects data samples directly from the original ensemble, avoiding the uncertainty induced from randomness when new samples are generated, thus more accurately preserving the original distribution. More importantly, using density plots to visualize the information of multivariate ensembles essentially requires estimations of individual bivariate PDFs, another computational and storage challenge. The sampling technique

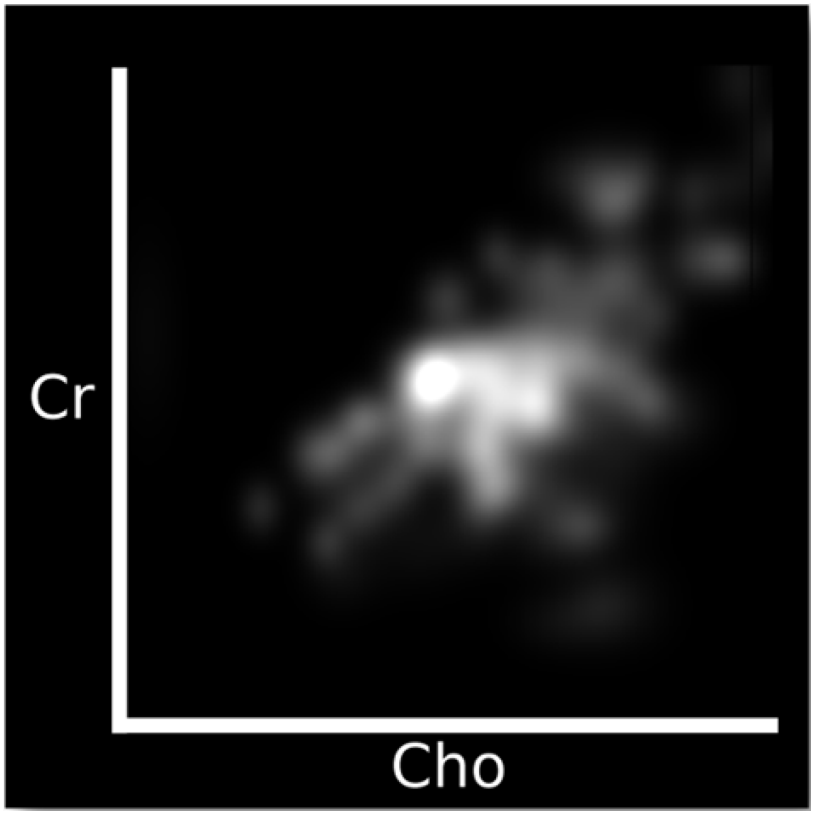
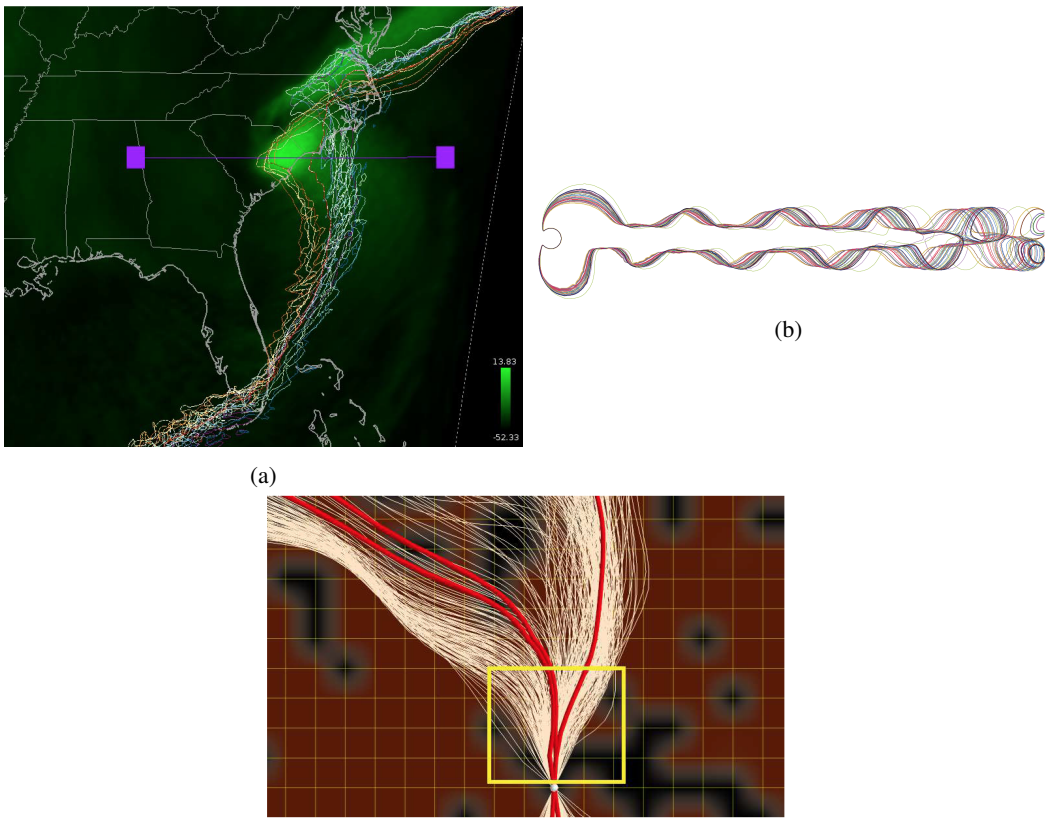



Figure 6.2: A density plot of two MR spectroscopy metabolites including choline and creatine, *Source*: [FKLT10].

developed here can potentially be generalized from 2D to higher dimensional spaces, meaning the members of this ensemble are abstracted as points in the higher dimensional space, and subsequently, the distributions of the variables included in the ensemble are preserved by a small subset of these higher dimensional points, simplifying the process of producing density plots.

A second example of how this research can possibly benefit the visualization community is by providing a new direction for depicting complex ensembles and, more importantly, their uncertainties. Currently, the most widely used approaches for visualizing ensembles are spaghetti plots and summary displays, both of which are usually integrated into ensemble visualization applications as they are complementary to each other: while spaghetti plots have the advantages of visually depicting as much information as possible, they potentially include visual clutter; on the other hand, while summary displays avoid the visual clutter, they do not represent the complete information of the ensembles, resulting in the difficulty of exploring and comparing the data. For instance, the Anscombes quartet [Ans73] contains four widely varied sets of numbers, the primary statistical summaries of which, including means, variances, correlations and regression lines, are, however, identical; thus, visualizing these different sets by displaying these statistical summaries produces identical visualizations. The direct ensemble visualization developed in this research appears as a hybrid of these two techniques, meaning that the explicitly portrayed representative collection of actual ensemble members are summarized based on the significant statistical features of the complete ensemble, thus keeping the advantages and minimizing the limitations of these two techniques. As the cognitive studies conducted in this research in collaboration with other researchers have supported the effectiveness of direct ensemble visualizations in the context of hurricane forecasting, a meaningful future direction is to extend this technique to various numerical simulation ensembles such as those of the water-vapor, temperature, and pressure of the 13 March 1993 "Superstorm" generated using the Weather Research and Forecast model [SZD+10] (Figure 6.3(a)); those of fluid flow simulation generated by perturbing the initial conditions [WMK13] (Figure 6.3(b)); and those of streamlines computed from vector fields [Hol15] (Figure 6.3(c)).

As these simulation ensembles are usually multimodal, the challenge in visualizing them is to preserving their modalities. To illustrate the possibility of applying the representative sampling technique of tracks to these ensembles, this technique is applied to a synthetic dataset containing two peaks of distributions is generated as shown in Figure 6.4(a). The result shown in Figure 6.4(b) where a small collection of tracks are selected, accurately achieves the two modalities of distributions included in the original dataset, supporting the potential of the technique introduced here to be extended to more general ensembles and, thus, supporting a wider use of direct ensemble visualizations. If these visualizations exhibit consistent effec-

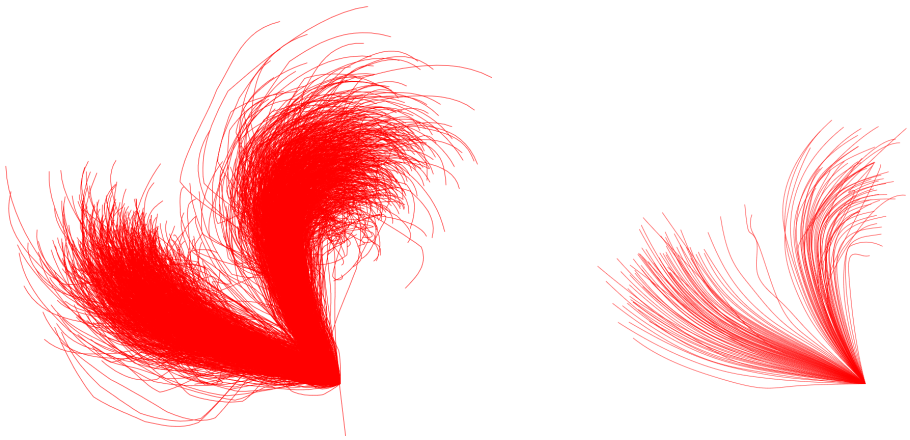



(a)

(b)

(c)

Figure 6.3: Examples of numerical simulation ensembles: (a) shows a temperature ensemble of the 13 March 1993 "Superstorm", *Source*: [SZD+10]; (b) shows an ensemble of isocontours of a pressure field generated from a fluid flow simulation, *Source*: [WMK13]; (c) shows an ensemble of streamlines integrated from a velocity field of a weather forecast, *Source*: [Hol15].



(a) (b)

Figure 6.4: A demonstration of using the sampling framework developed in this research to a track ensemble including a multimodal distribution: (a) is a plot of the ensemble, (b) shows a subset of representative tracks extracted from the ensemble.

tiveness as reported in the context of hurricane forecasting, this research would become an asset for a wide spectrum of currently limited visualization techniques.

# Bibliography


[ALP14] Ken Anjyo, John P Lewis, and Frédéric Pighin. Scattered data interpolation for computer graphics. In *ACM SIGGRAPH 2014 Courses*, page 27. ACM, 2014.

[Alt92] Naomi S Altman. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185, 1992.

[Ans73] Francis J Anscombe. Graphs in statistical analysis. *The American Statistician*, 27(1):17–21, 1973.

[Atg99] Frédéric Atger. Tubing: An alternative to clustering for the classification of ensemble forecasts. *Wea. Forecasting Weather and Forecasting*, pages 741–757, 1999.

[BHMG14] T. Barrett, M. Hegarty, G. McKenzie, and M. Goodchild. Am I really there? Evaluating visualizations of geospatial uncertainty. *Spatial Cognition poster*, September 2014.

[BL88] D. S. Broomhead and D. Lowe. Multivariable functional interpolation and adaptive networks. *Complex Systems 2*, pages 321–355, 1988.

[BLWS07] Kenneth Broad, Anthony Leiserowitz, Jessica Weinkle, and Marissa Steketee. Misinterpretations of the "cone of uncertainty" in Florida during the 2004 hurricane season. *Bulletin of the American Meteorological Society*, 88(5):651–667, 2014/11/11 2007.

[BMBW15] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

[Bro04] Ross Brown. Animated visual vibrations as an uncertainty visualisation technique. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 84–89. ACM, 2004.

[BS11] D.J. Bauer and S.K. Sterba. Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation. *Psychol Methods*, 16:373–90, 2011.

[BT07] D. Borland and R.M. Taylor. Rainbow color map (still) considered harmful. *Computer Graphics and Applications, IEEE*, 27(2):14–17, March 2007.

[CBC+01] J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, and T. R. Evans. Reconstruction and representation of 3d objects with radial basis functions. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 67–76, 2001.

[CCA96] S Chen, E.S. Chng, and K. Alkadhimi. Regularized orthogonal least squares algorithm for constructing radial basis function networks. *International Journal of Control*, 64(5):829–837, 1996.

[CCG91] S Chen, C.F.N. Cowan, and P.M. Grant. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on Neural Networks*, 2(2):302–309, 1991.

[CCM+14] Haidong Chen, Wei Chen, Honghui Mei, Zhiqi Liu, Kun Zhou, Weifeng Chen, Wentao Gu, and Kwan-Liu Ma. Visual abstraction and exploration of multi-class scatterplots. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):1683–1692, 2014.

[CCWA13] Jacob Cohen, Patricia Cohen, Stephen G West, and Leona S Aiken. *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge, 2013.

[CG15] Michael Correll and Michael Gieicher. Implicit uncertainty visualization: Aligning perception and statistics. *IEEE VIS 2015 Workshop "Visualization for decision making under uncertainty"*, 2015.

[CH13] J. Cox and D. House. Visualizing uncertainty as an interactive ensemble. *COSIT Workshop on Visually-Supported Reasoning with Uncertainty, accepted for presentation*, 2013.

[CHL13] J. Cox, D. House, and M. Lindell. Visualizing uncertainty in predicted hurricane tracks. *International Journal of Uncertainty Quantification*, 3(2):143–156, 2013.

[Coo86] Robert L. Cook. Stochastic sampling in computer graphics. *ACM Trans. Graph.*, 5(1):51–72, January 1986.

[DB72] Carl De Boor. On calculating with b-splines. *Journal of Approximation Theory*, 6(1):50–62, 1972.

[DH06] D. Dunbar and G. Humphreys. A spatial data structure for fast Poisson-disk sample generation. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*, 25(3):503–508, 2006.

[DJK+09] Mark Demaria, John, A. Knaff, Richard Knabb, Chris Lauer, Charles R. Sampson, and Robert T. Demaria. A new method for estimating tropical cyclone wind speed probabilities. *Wea. Forecasting*, 24:1573–1591, 2009.

[DKB+13] Mark DeMaria, John A Knaff, Michael J Brennan, Daniel Brown, Richard D Knabb, Robert T DeMaria, Andrea Schumacher, Christopher A Lauer, David P Roberts, Charles R Sampson, et al. Improvements to the operational tropical cyclone wind speed probability model. *Weather and Forecasting*, 28(3):586–602, 2013.

[ESV12] Ehsan Elhamifar, Guillermo Sapiro, and Rene Vidal. Finding exemplars from pairwise dissimilarities via simultaneous sparse recovery. In *Advances in Neural Information Processing Systems*, pages 19–27, 2012.

[FBW16] Florian Ferstl, Kai Burger, and Rudiger Westermann. Streamline variability plots for characterizing the uncertainty in vector field ensembles. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):767–776, 2016.

[FKLT10] David Feng, Lester Kwock, Yueh Lee, and Russell Taylor. Matching visual saliency to confidence in plots of uncertain data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):980–989, 2010.

[Har03] Mark Harrower. Representing uncertainty: Does it help people make better decisions. In *UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop*, pages 18–20, 2003.

[Hol15] Brad Eric Hollister. *Visualizing multimodal uncertainty in ensemble vector fields*. University of California, Santa Cruz, 2015.

[HTER04] Christopher G. Healey, Laura Tateosian, James T. Enns, and Mark Remple. Perceptually based brush strokes for nonphotorealistic visualization. *ACM Trans. Graph.*, 23(1):64–96, January 2004.

[KSD+07] John A Knaff, Charles R Sampson, Mark DeMaria, Timothy P Marchok, James M Gross, and Colin J McAdie. Statistical tropical cyclone wind radii prediction using climatology and persistence. *Weather and Forecasting*, 22(4):781–791, 2007.

[Lan69] T Lang. Rules for robot draughtsmen. *Geographical Magazine*, 42(1):50–51, 1969.

[Lew05] J.M. Lewis. Roots of ensemble forecasting. *Montly Weather Rev*, pages 1865–1885, 2005.

[Lou14] National Weather Service Louisville. Weather forecast uncertainty, Jan 2014. `http://www.crh.noaa.gov/Image/lmk/pdf/WeatherForecastUncertainty.pdf`.

[LY90] Liu and Regina Y. On a notion of data depth based on random simplices. *The Annals of Statistics*, 18(1):405–414, 1990.

[MBOP13] Robert Meyer, Kenneth Broad, Ben Orlove, and Nada Petrovic. Dynamic simulation as an approach to understanding hurricane risk response: Insights from the stormview lab. *Risk Analysis*, 33(8), 2013.

[MBPP96] F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliagis. The ecmwf ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, 122(529):73–119, 1996.

[MF92] Michael McCool and Eugene Fiume. Hierarchical poisson disk sampling distributions. In *Proceedings of the conference on Graphics interface*, volume 92, pages 94–105, 1992.

[Mur88] JM Murphy. The impact of ensemble forecasts on predictability. *Quarterly Journal of the Royal Meteorological Society*, 114(480):463–493, 1988.

[MWK14] M. Mirzargar, R.T. Whitaker, and R.M. Kirby. Curve boxplot: Generalization of boxplot for ensembles of curves. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2654–2663, Dec 2014.

[NIS92] NIST. Statements of uncertainty associated with measurement results. Technical report, NIST Technical Communication Program, 1992.

[NOA99] "National Hurricane Center" NOAA. *NOAA's Hurricane Officials on Target with Pre-seasonal Storm Forecast; Close 1999 Season Reporting Above Average Fury And Floods*, 1999.

[NOA00] "National Hurricane Center" NOAA. Keith graphics archive, 2000. `http://www.nhc.noaa.gov/archive/2000/KEITH_graphics.html`.

[NOA09] "National Hurricane Center" NOAA. Technical summary of the national hurricane center track and intensity models, July 2009. `http://www.nhc.noaa.gov/pdf/model_summary_20090724.pdf`.

[NOA12] "National Hurricane Center" NOAA. Isaac graphics archive, 2012. `http://www.nhc.noaa.gov/archive/2012/graphics/al09/loop_3W.shtml`.

[NOA14] "National Hurricane Center" NOAA. Surface wind speed probabilities (120 hours), December 2014. `http://www.nhc.noaa.gov/archive/2012/ISAAC_graphics.shtml`.

[NOA15a] "National Hurricane Center" NOAA. Atlantic hurricane outlook, Aug 2015. `http://www.cpc.ncep.noaa.gov/products/outlooks/hurricane.shtml`.

[NOA15b] "National Hurricane Center" NOAA. National hurricane center forecast verification, March 2015. http://www.nhc.noaa.gov/verification/verify5.shtml.

[NOA17a] "National Hurricane Center" NOAA. Annual average official intensity errors for atlantic basin tropical cyclones for the period 1990-2016, 2017. http://www.nhc.noaa.gov/verification/pdfs/1990-present_OFCL_ATL_annual_int_errors.pdf.

[NOA17b] "National Hurricane Center" NOAA. Annual average official track errors for atlantic basin tropical storms and hurricanes for the period 1970-2016, 2017. http://www.nhc.noaa.gov/verification/pdfs/1970-present_OFCL_ATL_annual_trk_errors_noTDs.pdf.

[NOA17c] "National Hurricane Center" NOAA. National hurricane center forecast verification, 2017. http://www.nhc.noaa.gov/verification/verify4.shtml.

[NOA17d] "National Hurricane Center" NOAA. Saffir-simpson hurricane wind scale, 2017. http://www.nhc.noaa.gov/aboutsshws.php.

[O+96] Mark JL Orr et al. Introduction to radial basis function networks, 1996.

[Pan08] Alex Pang. Visualizing uncertainty in natural hazards. In *Risk Assessment, Modeling and Decision Support*, volume 14 of *Risk, Governance and Society*, pages 261–294. Springer Berlin Heidelberg, 2008.

[PH11] Kai Pöthkow and Hans-Christian Hege. Positional uncertainty of isocontours: Condition analysis and probabilistic measures. *Visualization and Computer Graphics, IEEE Transactions on*, 17(10):1393–1406, 2011.

[Pot10] Kristin Potter. *The visualization of uncertainty*. The University of Utah, 2010.

[PWL96] Alex T. Pang, Craig M. Wittenbrink, and Suresh K. Lodh. Approaches to uncertainty visualization. *The Visual Computer*, 13:370–390, 1996.

[RABCdT11] S.W. Raudenbush, Y.F. Cheong A.S. Bryk, R.T. Congdon, and M. du Toit. *Hierarchical Linear Modeling 7*, 2011.

[RB02] S.W. Raudenbush and A.S. Bryk. Hierarchical linear models: Applications and data analysis methods. *Sage*, 1, 2002.

[RBP+16] Ian T Ruginski, Alexander P Boone, Lace M Padilla, Le Liu, Nahal Heydari, Heidi S Kramer, Mary Hegarty, William B Thompson, Donald H House, and Sarah H Creem-Regehr. Non-expert interpretations of hurricane forecast uncertainty visualizations. *Spatial Cognition & Computation*, 16(2):154–172, 2016.

[RR96] Peter J. Rousseeuw and Ida Ruts. Algorithm AS 307: Bivariate location depth. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(4):pp. 516–526, 1996.

[RVC02] L. Rocha, L. Velho, and P.C.P. Carvalho. Image moments-based structuring and tracking of objects. In *Computer Graphics and Image Processing, 2002. Proceedings. XV Brazilian Symposium on*, pages 99–105, 2002.

[SC06] Wenzhong Shi and ChuiKwan Cheung. Performance evaluation of line simplification algorithms for vector generalization. *The Cartographic Journal*, 43(1):27–44, 2006.

[Sch46] Isaac J Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions. part a. on the problem of smoothing or graduation. a first class of analytic approximation formulae. *Quarterly of Applied Mathematics*, 4(1):45–99, 1946.

[SDR00] David B. Stephenson and Francisco J. Doblas-Reyes. Statistical methods for interpreting monte carlo ensemble forecasts. *Tellus A*, 52(3):300–322, 2000.

[SEZ12] W. Schneider, A. Eschman, and A Zuccolotto. *E-Prime user's guide*, 2012.

[Ste09] B. Stearns. *The Homeowner's Hurricane Handbook*. Skyhorse Pub., 2009.

[SZD+10] Jibonananda Sanyal, Song Zhang, Jamie Dyer, Andrew Mercer, Philip Amburn, and Robert Moorhead. Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1421–1430, 2010.

[TK93] Zoltan Toth and Eugenia Kalnay. Ensemble forecasting at nmc: The generation of perturbations. *Bulletin of the American Meteorological Society*, 74:2317–2330, 1993.

[TK97] Zoltan Toth and Eugenia Kalnay. Ensemble forecasting at ncep and the breeding method. *Monthly Weather Review*, 125:3297–3319, 1997.

[War04] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.

[Wei10] Li-Yi Wei. Multi-class blue noise sampling. *ACM Transactions on Graphics (TOG)*, 29(4):79, 2010.

[WMK13] R.T. Whitaker, M. Mirzargar, and R.M. Kirby. Contour boxplots: A method for characterizing uncertainty in feature sets from simulation ensembles. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2713–2722, Dec 2013.

[WPL96] C.M. Wittenbrink, A.T. Pang, and S.K. Lodha. Glyphs for visualizing uncertainty in vector fields. *Visualization and Computer Graphics, IEEE Transactions on*, 2(3):266–279, Sep 1996.

[XHWR06] Zaixian Xie, Shiping Huang, Matthew O Ward, and Elke A Rundensteiner. Exploratory visualization of multivariate data with variable quality. In *Visual Analytics Science And Technology, 2006 IEEE Symposium On*, pages 183–190. IEEE, 2006.

[Yuk15] C. Yuksel. Sample elimination for generating Poisson disk sample sets. *Computer Graphics Forum (Proceedings of Eurographics 2015)*, 34(2):25–32, 2015.